

VU Research Portal

Verkenning van de toepassingsmogelijkheid van floating car data voor het beschrijven van verplaatsingsgedrag

Koomen, E.; Andree, B.P.J.; Jacobs-Crisioni, C.G.W.

2018

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Koomen, E., Andree, B. P. J., & Jacobs-Crisioni, C. G. W. (2018). *Verkenning van de toepassingsmogelijkheid van floating car data voor het beschrijven van verplaatsingsgedrag*. Vrije Universiteit Amsterdam.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

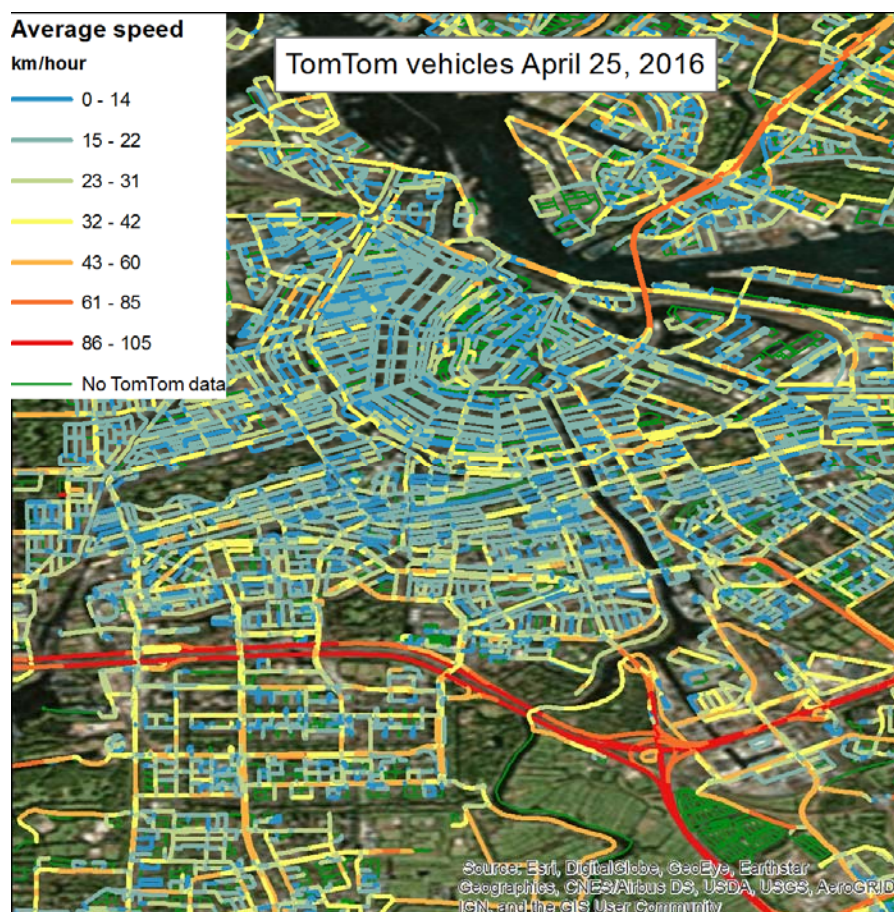
Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Verkenning van de toepassingsmogelijkheden van floating car data voor het beschrijven van verplaatsingsgedrag



Eric Koomen, Bo Andrée, Chris Jacobs-Crisioni
Maart, 2018



Opdracht voor Rijkswaterstaat - Centrale Informatievoorziening

In het kader van de raamovereenkomst Samenwerking en Kennisuitwisseling op het gebied van Ruimtelijke Informatievoorziening

Contact informatie:
Vrije Universiteit Amsterdam
School of Business and Economics
Department of Spatial Economics/ Spatial Information Laboratory (SPIN/*lab*)
De Boelelaan 1105
1081 HV Amsterdam
Netherlands
Phone: +31 20 5986095
Email: e.koomen@vu.nl
Website: <http://spinlab.vu.nl/>

Samenvatting

RWS en Vrije Universiteit Amsterdam zijn op het gebied van Ruimtelijke Informatievoorziening een kennisuitwisselings- en onderzoeksrelatie aangegaan voor een periode van 4 jaar. Deze relatie is april 2015 vastgelegd in de Raamovereenkomst Samenwerking en Kennisuitwisseling op het gebied van Ruimtelijke Informatievoorziening. Het voorliggende onderzoeksrapport is opgesteld als onderdeel van deze raamovereenkomst. Het project is inhoudelijk gedefinieerd en uitgevoerd in overleg met Frank Hofman, Dusica Krstic-Joksimovic en Marcel Mulder van afdeling Modellen en Applicaties van het onderdeel Water, Verkeer en Leefomgeving (WVL) van Rijkswaterstaat.

Doel van dit project is na te gaan in hoeverre zogeheten floating car data kan bijdragen aan de beschrijving van verplaatsingsgedrag. Deze term beschrijft een bredere set aan voertuig gebaseerde verplaatsingsdata. Naast data afkomstig van navigatiesoftware (zoals in-car navigation data van bijvoorbeeld TomTom) betreft dit bijvoorbeeld data van andere diensten (zoals file-informatie) of wagenparkbeheerders die hun auto's volgen. We hebben ons daarbij allereerst gericht op de vraag of de beschikbare data behulpzaam kunnen zijn bij het vaststellen van zogenaamde herkomst-bestemmingsmatrices die gebruikt worden in de verkeersmodellen zoals WVL die toepast. Om deze vraag te beantwoorden hebben we een literatuurinventarisatie gedaan en de beschikbare data van een belangrijke in-car navigation data leverancier (TomTom) geanalyseerd.

De belangrijkste conclusie van deze verkenning is dat de beschikbare data die verkeersintensiteiten en snelheden per wegvak beschrijft, niet geschikt is voor het opstellen van herkomst-bestemmingsmatrices. Deze wegsegment data is wel bruikbaar voor een groot aantal andere toepassingen om verplaatsingsgedrag te analyseren. Om dat te illustreren is een analyse gedaan van reistijd-betrouwbaarheid in de ochtendspits van twee specifieke deur-tot-deur trajecten. Hierbij is onderscheid gemaakt in de delen die over het hoofdwegennet en het onderliggend wegennet zijn afgelegd. Dergelijke deur-tot-deur analyses van reistijden zijn minder goed mogelijk met de tot nu toe gebruikte verkeersgegevens omdat er minder informatie beschikbaar was over de reistijden op het onderliggend wegennet. Het analyseren van trajecten is namelijk niet mogelijk op basis van de punt-informatie uit bijvoorbeeld tellussen omdat die bron vooral voor de drukke autosnelwegen beschikbaar zijn. Floating car data kan hierbij dus van meerwaarde zijn om reistijdbetrouwbaarheidanalyses te doen voor specifieke trajecten die bestaan uit delen op het hoofdwegennet en het onderliggend wegennet op basis van een groot aantal waarnemingen.

Voor deze analyse is gebruik gemaakt van floating-car data die is verzameld door BeMobile en die mede door Rijkswaterstaat is aangekocht. Deze data is door de Nationale databank wegverkeersgegevens (NDW) beschikbaar gemaakt voor dit onderzoek. Deze rapportage sluit af met een beknopte beschrijving van mogelijke vervolgt toepassingen en geeft tevens aan welke andere floating car data sets beschikbaar zijn voor het analyseren van het verplaatsingsgedrag van groepen van individuen.

Deze studie was niet mogelijk geweest zonder de medewerking van de verschillende floating car data leveranciers. Specifiek willen we Daan Scheer (TomTom), Bas Jansen Venneboer (Local-eyes) en Marthe Uenk-Telgen (NDW) bedanken voor hun tijd en de gegevens die ze beschikbaar stelden. Daarnaast zijn wij onze VU-collega's Paul Koster en Erik Verhoef erkentelijk voor hun suggesties voor de analyse van de gebruikte gegevens.

Inhoudsopgave

Samenvatting	3
1. Introductie.....	7
2. Literature review	8
3. Dataverkenning en analyse	11
3.1. Data visualisatie	11
3.2. Analyse van reistijdbetrouwbaarheid in de ochtendspits	15
4. Conclusie	24
4.1. Recent onderzoek naar bruikbaarheid floating car data per weg segment	24
4.2. Mogelijk vervolgonderzoek met floating car data op wegsegment niveau	26
4.3. Alternatieve floating car data	29
Referenties.....	30

1. Introductie

In de besluitvorming over investeringen in het rijkswegennet spelen verkeers- en vervoers-modellen een belangrijke rol. De strategische verkeer- en vervoermodellen Nederlands Regionaal Model (NRM) en Landelijk Model Systeem (LMS) van het Ministerie van Infrastructuur en Waterstaat worden gebruikt voor het opstellen van prognoses van mobiliteitsontwikkelingen, van de belastingen van het hoofdwegennet en het spoornetwerk. Op basis van deze prognoses worden de effecten van beleidsmaatregelen voorspeld en worden effecten op de leefomgeving berekend. Dergelijke modellen worden gekalibreerd o.a. op basis van gemeten verkeersintensiteiten en enquêtegegevens uit het Onderzoek Verplaatsingen in Nederland (OVIN, voorheen Mobiliteitsonderzoek Nederland, MON, zie CBS en RWS, 2016). Op basis van dit onderzoek wordt voor verschillende persoonskenmerken/ huishoudensgroepen het gemiddelde verplaatsingsgedrag beschreven, ook voor werkdagen in de vorm van regionale herkomst- en bestemmingsmatrices.

De toenemende beschikbaarheid van nieuwe (big) databronnen en data sciences methoden opent mogelijkheden om op een alternatieve manier verplaatsingsgedrag te analyseren en te gebruiken bij het genereren van herkomst-bestemmingsmatrices. De afdeling Modellen en Applicaties van het onderdeel Water, Verkeer en Leefomgeving (WVL) van Rijkswaterstaat verkent momenteel de mogelijkheden van mobiele telefoondata om verplaatsingsgedrag te analyseren en is zeer geïnteresseerd in de mogelijkheden die andere databronnen op dit vlak kunnen bieden. Om in deze interesse te voorzien zal dit onderzoeksproject verkennen in hoeverre het mogelijk is floating car data te gebruiken om verplaatsingsgedrag te beschrijven. Het uiteindelijke doel is die informatie te gebruiken bij het actualiseren van de verkeer- en vervoermodellen van RWS. Floating car data is een verzamelnaam voor data die de verplaatsingen van individuele voertuigen beschrijft en die bijvoorbeeld wordt ingewonnen door navigatiesoftware in voertuigen (in-car navigation data) of via apps op de mobiele telefoon van weggebruikers die behulpzaam zijn bij routekeuze (zoals GoogleMaps) of waarschuwen voor files of flitspalen (zoals Flitsmeister).

Deze beknopte studie is opgezet en uitgevoerd in nauwe samenwerking met het onderdeel Water, Verkeer en Leefomgeving (WVL) van Rijkswaterstaat (meer specifiek: Frank Hofman, Dusica Krstic-Joksimovic en Marcel Mulder). Tussentijdse resultaten zijn met hen gedeeld en veranderingen in de oorspronkelijke onderzoeksopzet zijn met hen besproken.

Doelstelling en eindresultaat

Centraal in dit project staat de vraag: in hoeverre kan floating car data bijdragen aan een betere beschrijving van verplaatsingsgedrag in termen van: herkomst en bestemming; verplaatsingsmotief en routekeuze? In het bijzonder is hierbij gekeken hoe bruikbaar deze gegevens zijn voor het uitvoeren van analyses naar de betrouwbaarheid van reistijden.

Deze studie start met een beknopte literatuurinventarisatie van reeds bestaande toepassingen van floating car data voor het beschrijven van verplaatsingsgedrag. Dit deel is opgesteld in het Engels om onderdeel te kunnen vormen van een wetenschappelijk artikel. Vervolgens is een verkenning uitgevoerd met een selectie van floating car data voor een stedelijke regio in Nederland om te achterhalen welke data voor onderzoek beschikbaar is en welke mogelijkheden deze gegevens bieden om verplaatsingsgedrag te beschrijven. In deze opstartfase van het project is vooral kennis opgedaan hoe dergelijke data benaderd en gevisualiseerd kan worden en in hoeverre het mogelijk is herkomst, bestemming en doel af te leiden uit de beschikbare verplaatsingsdata. In aansluiting daarop is een beknopte studie gedaan naar een specifiek aspect van verplaatsingsgedrag dat met floating car data kan worden onderzocht: de betrouwbaarheid van reistijden. Afsluitend beschrijven we aanvullende onderzoeksmogelijkheden met de beschikbare data en geven we een beknopte reflectie op verkeerskundig onderzoek dat mogelijk is met floating car data.

2. Literature review

This section documents some key findings from literature on using floating car data in establishing Origin-Destination (OD) matrices. Concepts that are of importance when working with the particular Tom-Tom segment dataset that has been made available for the initial phase of this research project are highlighted in footnotes.

The following general strategies can be categorized that estimate traffic flows between sets of origins and destinations:

- 1) statistical inference of the OD matrix from observed partial network flows;
- 2) statistical inference of the OD matrix from observed total in and out flows of the defined regions;
- 3) calibration of an existing OD matrix based on a target matrix derived from observed flows on the network and possibly a set of constraints; and
- 4) direct extraction of origins and destinations from georeferenced trip data generated by Floating Car Technologies.

The four strategies rely on different aggregations of traffic data. The first strategy relies on observed flows aggregated at the segment level, whereas the second strategy relies on total in and out flows aggregated at a regional level. The third strategy relies on a target matrix extracted from any dataset to update the matrix currently in use, and strategy four relies on the individual, non-aggregated, trip data that is aggregated directly into the OD matrix format. The latter seems most powerful, but potentially also raises privacy concerns. Therefore data providers may be reluctant to share such data. An important, general issue in dealing with the floating car data of any data provider relates to weighting and expansion of the observations to their set of sampled individuals to the total number travelers.

The first method should be feasible with the segment data we can get access to, but bias of the Tom-Tom data could play a disastrous role. The second method is widespread, and estimation of the gravity model is the most well-known method. The third method seems very artificial, and the product might be more strongly generated by assumptions than by observations, but it is used often in transport modeling. This method also depends on quality input that might not be available. It can be used to update old OD matrices that are generated from obsolete trip diaries. The update method may be of importance for dynamic OD matrices, or in situations where there are strong reasons to control the rate of change between old and new OD matrices, for example when it is known that data from a particular period was strongly dependent on (macro)-economic factors. The fourth method is a direct stocktaking method of starting and ending points of recorded trips. If the data covers all regions, the resulting OD matrix is complete. However, if there are missing values at locations where trips would be expected, these would have to be fitted in by using historic data or statistical methods, for example by fitting a gravity model and predicting the fitted model, or, as I would theorize, by predicting NA values by spatial interpolation.

The following subsections briefly summarise the findings of several relevant resources. More general introductions to Origin-Destination matrix estimation can be found elsewhere (Bera & Rao, 2011; Simoni, 2014).

Building practical Origin-Destination (od/trip) matrices from automatically collected GPS data

This paper by Van Vuren et al. (2011) discusses the pros and cons of GPS data that includes time intervals with locational point data. Origin and destination counts are obtained at an aggregated level through additional data on ignition on/off data. Focus is in comparison with conventional collection methods, and the benefits of GPS methods. Locational errors of GPS origins are addressed,

and the paper proposes linking origins to the last known destination as a potential way of reducing the error¹. The paper is fairly introductory and has a GIS focus.

Dynamic Origin-Destination matrix estimation from traffic counts and automated vehicle identification data

Classic paper by Van der Zijpp (1997) that deals with statistical issues of OD estimation from vehicle identification/recognition data. An important issue that is discussed is that OD estimation is an underspecified problem, multiple OD matrices can generate the same observations on the network.² The paper proposes a method that restricts the time-varying rate of change in split probabilities as an identification strategy. Central to the paper is measurement errors in the identification rate, and the focus is to explore methods like the Kalman filter, and its Bayesian interpretation, to filter out the unobserved time-varying split probabilities.³ A key requirement is that registration stations correspond to network entries and exits.⁴

The Origin–Destination matrix estimation problem - analysis and computations

PhD thesis of Anders Peterson (2007) that starts with a useful introduction on the traffic assignment problem as a background to the OD estimation problem.⁵ The focus is on dealing with congestion in the estimation of latent locational traffic demand.⁶ Generally, the manuscript makes a few good points, but it focusses strongly on the algorithms. The use of these algorithms is questionable as they assume a suitable target OD matrix is available. Since the real OD matrix is unknown, the reliability of the developed methods might be difficult to assess.

Dynamic OD matrix estimation using floating car data

A very relevant and complete MSc thesis written by Rúna Ásmundsdóttir (2008) at and for Rijkswaterstaat. The report focusses on estimation of the OD matrix using Floating Car Data, with a specific focus on Taxi data. The main algorithm suggested initiates with an a priori matrix that is

¹ This could be a potential bias, but there is no remedy as the data is already aggregated.

² This could be true if not all locations of origins and destinations are observed - which is necessarily the case if only data from registration points that do not correspond to origins is available - vehicles entering at one point in the network could have come from different unobserved origins. In that case the system is under identified – the number of registration points is smaller than the number of origin destination combinations - but that could be solved by looking at higher aggregation levels such that the number of equations in the true OD generating process correspond to the number of modelled ones. Identifiability and uniqueness of the OD matrix is a separate problem, a proof of uniqueness will not be discussed as it would vary among estimation specifications. The problem is documented, and motivates the use of a target matrix.

³ Measurement errors in the time dimension are also a problem with GPS segment data as the share of Tom-Tom devices might vary randomly over time and is not observed. Kalman smoothers or polynomial smoothers might be important ways to filter out these measurement errors, if the goal is to estimate time-varying metrics from the data.

⁴ Segment intensity data might indirectly provide a (normalized) measure of entries and exits at the node level and thus an indication of the total flows into origins and destination.

⁵ Could be useful to provide context to the problem at hand. OD estimation is the inverse of traffic assignment. With segment data, we might observe a strong correlate of the true flows, but a challenging aspect is the handling of measurement errors and consistency problems.

⁶ Observed OD traffic is not completely in par with the underlying OD demand, as demand and congestion endogenously interact, such that the observed outcome is biased with respect to the true unobserved demand. A potentially interesting idea is that during uncongested times, the observed flows are easily linked to the underlying OD demand. This could provide an initial target matrix, and the time-varying matrices could iteratively be estimated based on using the previous matrix as a target matrix. This deals both with identifiability and the congestion bias.

assigned to a network and iteratively adjusted till generated flows match observed flows. The main suggestion is that taxi data can be used to derive the a priori matrix. A review of floating car data applications is provided, different methods for OD estimation are reviewed, the data and its pre-processing are discussed. The thesis concludes that a-priori matrix estimation is feasible, but taxi data is not representative enough to solely rely on.⁷

Estimating Origin-Destination flows using opportunistically collected mobile phone location data from one million users in Boston Metropolitan Area

This paper by Calabrese and others (2011) carefully documents the estimation of an OD matrix from mobile phone data in the Boston MSA. The paper begins with deriving a number of basic statistics such as the distribution of trip length and the average number of trips in weekdays and weekends.⁸ A comparison is made with US National Household Travel Survey to evaluate sample bias.⁹ Cell-phone usage density is compared to population density to evaluate coverage and locational representativeness. The strategy of converting locational space-time point data to time sequence of OD matrices is well-documented and applicable to other floating sensors.¹⁰ The paper concludes with an overview of limitations of mobile phone data for OD estimation.¹¹

Integration of traffic flow relations in a simulation environment

A research paper by Reinthaler and Zajicek (2009) that describes a method to find OD relations useful for a microscopic traffic model. The method is adopted and empirically tested in a case study with taxi data. The paper starts with a concise and clear summary of transport modelling and the role of the OD matrix in the endeavour. The data description is clear, but the methods are vaguely described without any equations. This makes it difficult to reproduce the method.

Some comments on Origin–Destination matrix estimation

Another classic paper by Hazelton (2003) that details statistical estimation of the OD matrix from link counts. The introduction contains a good example to understand why counts at the segment level, are insufficient to estimate the underlying OD matrix. The straightforward problem is that link count observations are not the same as route choice observations, multiple OD matrices – essentially the counts per possible route - can generate the same link count observations.¹² The problem is related to invertibility, one can calculate the mean of a set of real numbers but if only the mean is known, the possible sets that have the same mean is infinite. A variety of statistical techniques that exploit secondary statistical properties of the data are discussed.

⁷ Tom-Tom should be more representative, so it should be useful for estimating an a priori matrix that is than further optimized with iterative algorithms.

⁸ A useful starting point.

⁹ A recommended strategy if such a dataset, or better if its statistics, are available.

¹⁰ That would be the way to go for floating car data, and the data setup should be put forward as the required one.

¹¹ Most remarks should be evaluated for the Tom-Tom data as well.

¹² The introductory example is a perfect starting point for this research. The ultimate added value of floating car sensors is that we do in fact observe flow counts instead of link counts, if data on the individual trips is available.

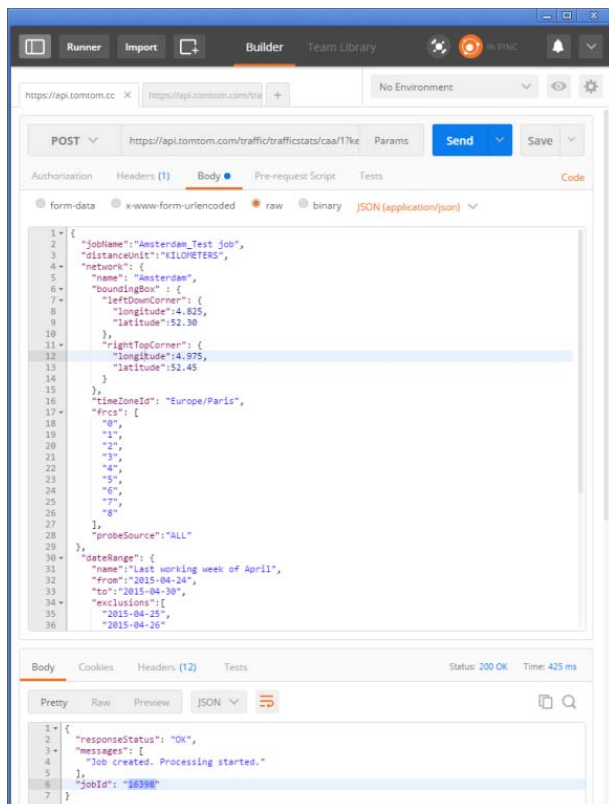
3. Dataverkenning en analyse

In het toegepaste deel van dit project hebben we een verkenning gedaan met verkeersgegevens afkomstig van floating car data bronnen. We gebruiken hiervoor data van TomTom en BeMobile. De TomTom data is toegepast in een eerste visuele verkenning uitgevoerd in 2016. Vanaf 2017 beheert NDW floating car data die van BeMobile is aangeschaft. Met deze data is een analyse van reistijdbetrouwbaarheid in de ochtendspits gedaan.

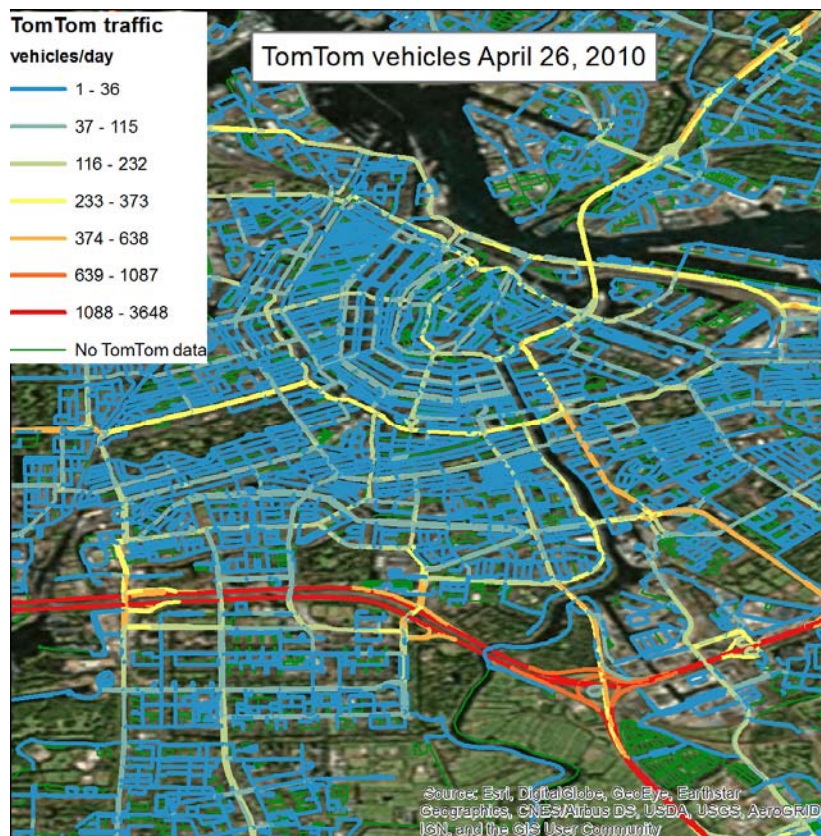
3.1. Data visualisatie

Voor dit project is TomTom zo vriendelijk geweest ons voor een beperkte periode toegang te verlenen tot hun Developer portal. Middels een application programming interface (API, zie Figuur 1) kan een enorme geodatabase met on-line traffic data ontsloten worden. Deze bevat onder meer: real-time flows en incidents; en “Historische” trafficstats. Die laatste groep heeft onze interesse en geeft informatie over (variatie in) snelheid en intensiteit op basis van het aantal TomTom gebruikers.

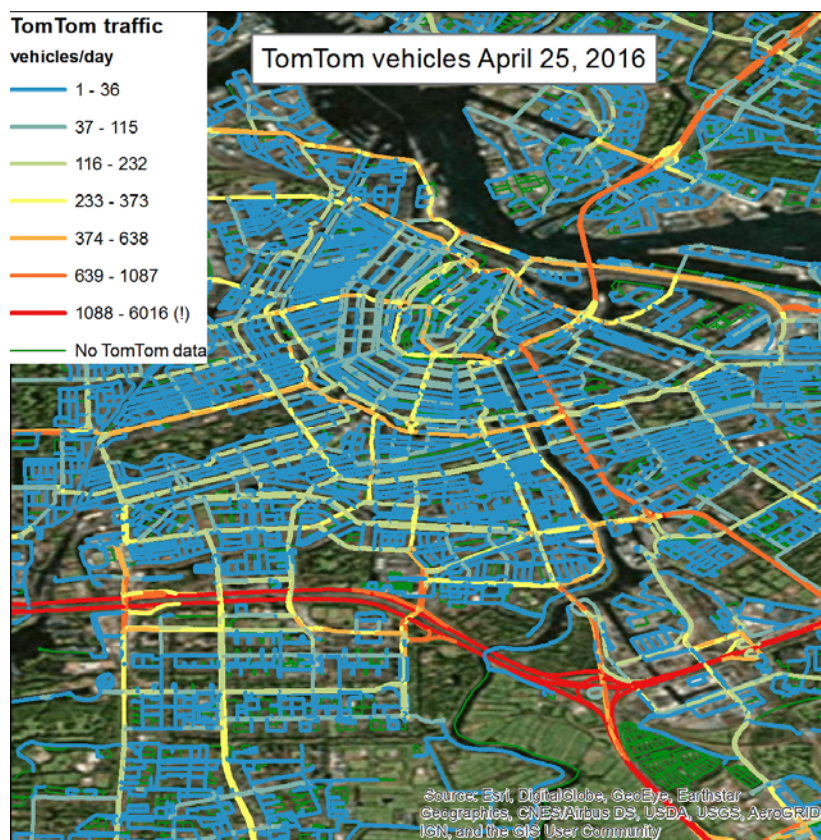
De trafficstats data is opgeslagen per wegsegment, rijrichting, dag (1/365) en tijdstip. Voor de geometrie van het netwerk is gebruik gemaakt van shapefiles van teleatlas. De data is op te vragen per route of gebied met een maximum van respectievelijk 200 km of 20.000 km². Voor Nederland is de data beschikbaar vanaf 2008. Daarnaast zijn er gegevens beschikbaar voor nog 66 andere landen. Het is mogelijk de data op te vragen voor een bepaald moment op een bepaalde dag op een specifieke plaats, of als gemiddelde over een bepaald tijdvak en periode. In onze verkenning hebben we data opgevraagd voor de regio Amsterdam met in totaal 50.560 weg-segmenten. De respons tijd (wachtijd totdat data downloadbaar staat) wisselde hierbij tussen de 15 minuten en 12 uur. Figuur 2 tot Figuur 4 geven voorbeelden van de opgevraagde data voor twee willekeurige maandagen: een in 2010 en 2016.



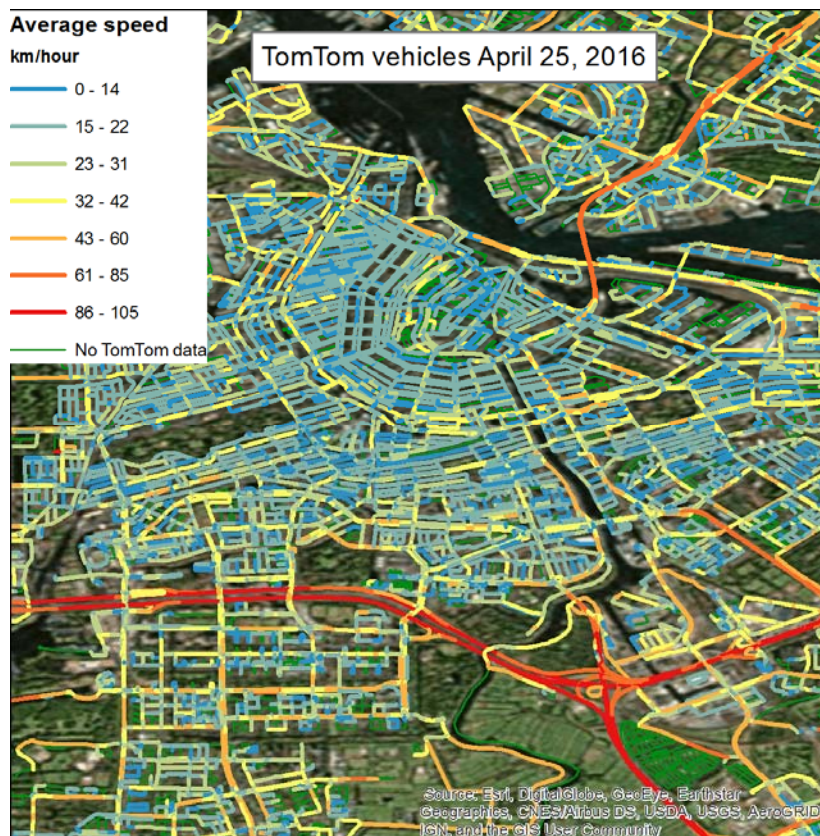
Figuur 1 Toegang krijgen tot TomTom data via een API



Figuur 2 Visualisatie data willekeurige maandag in 2010



Figuur 3 Visualisatie data willekeurige maandag 6 jaar later



Figuur 4 Naast intensiteit kan ook snelheid weergegeven worden

Het grote voordeel van deze data boven klassieke verkeerstelgegevens is dat ze niet gebonden zijn aan een beperkt aantal meetpunten. Omdat er veel TomTom gebruikers zijn en deze gebruik kunnen maken van alle wegen, bestrijkt deze data in principe het hele gehele wegennetwerk. Dat geldt zeker als data over langere tijdsperiode wordt opgevraagd. Deze data geeft dan ook een behoorlijk gedetailleerd inzicht in het gebruik van het onderliggend wegennet waarover Rijkswaterstaat zelf weinig gegevens heeft. Via de TomTom portal zijn geen gegevens op te vragen over individuele trips.

Op basis van deze gegevens is in het in principe ook mogelijk om veranderingen door de tijd heen te analyseren. Tabel 1 geeft een voorbeeld van een dergelijk vergelijk en laat zien dat voor deze twee willekeurige maandagen de gemiddelde snelheid is afgenomen (van 29 naar 27 km/u) terwijl de intensiteit fors toenam. Dergelijke analyses zullen voor veranderingen in gemiddelde snelheid (doorstroming) redelijk betrouwbaar zijn. Maar voor veranderingen in intensiteit is dat uiteraard sterk afhankelijk van het marktaandeel TomTom en het relatieve aantal gebruikers van hun applicatie. Voor het getoonde verschil tussen deze twee willekeurige dagen geldt natuurlijk dat dit beïnvloed kan zijn door een groot aantal factoren zoals bijvoorbeeld weersomstandigheden.

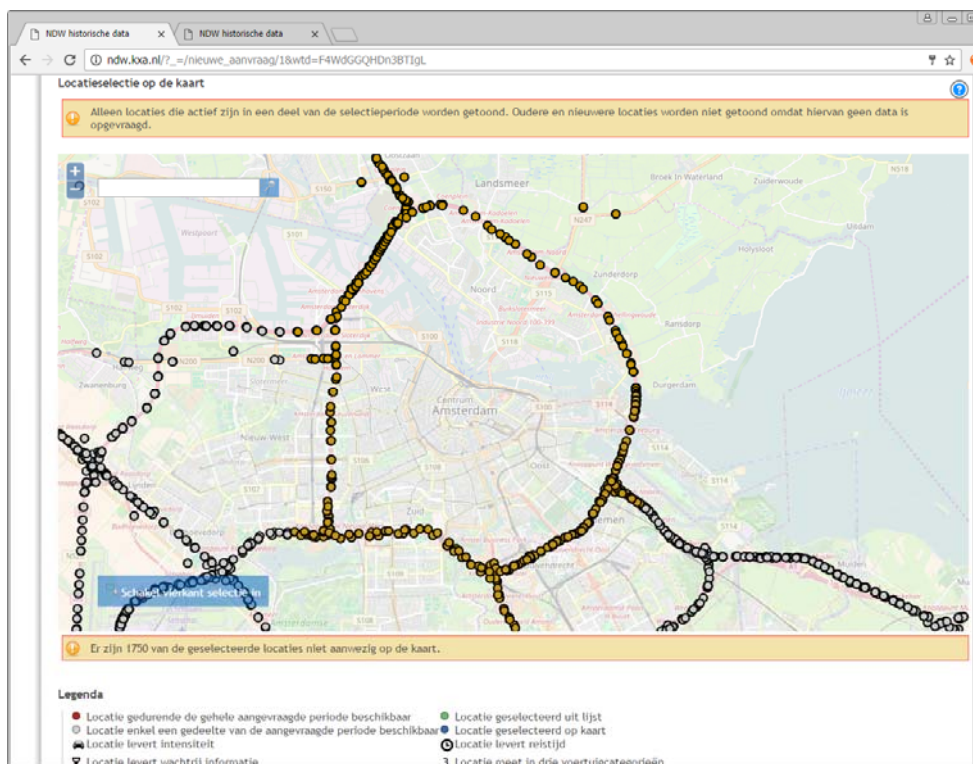
Tabel 1 Voorbeeld van daggemiddelde waarden voor snelheid (harmonisch) en intensiteit in de opgevraagde gegevens voor de regio Amsterdam zoals getoond in voorgaande figuren.

	N	Minimum	Maximum	Gemiddelde	Standaard afwijking	Gewogen gemiddelde*
26 april 2010						
Lengte (m)	34423	0.76	1935.6	59.7	71.8	
Intensiteit (aantal TomTom'ers)	34423	1	3648	95.7	334.2	148.7
Gemiddelde snelheid (km/u)	34423	0.19	142.0	24.1	16.6	29.0
25 april 2016						
Lengte (m)	37157	0.76	1935.6	60.3	74.0	
Intensiteit (aantal TomTom'ers)	37157	1	6016	150.0	505.2	225.9
Gemiddelde snelheid (km/u)	37157	0.13	153.0	22.5	15.1	27.3

* De daggemiddelde snelheid per wegvak is berekend als harmonisch gemiddeld zonder rekening te houden met verschillen in intensiteit. De gemiddelde snelheid over de wegvakken (met elk een verschillende lengte) is in de laatste kolom gewogen voor weglengte.

Representativiteit TomTom gegevens

De snelheid in deze dataset lijkt een goede maatstaf voor de gemiddelde snelheid op een wegvak omdat het niet waarschijnlijk is dat TomTom-gebruikers veel sneller of langzamer zullen rijden dan anderen. Om op basis van het aantal TomTom gebruikers op een wegvak iets te zeggen over het totaal aantal weggebruikers zijn aanvullende gegevens nodig. Door gebruik te maken van gegevens van bijvoorbeeld verkeerstellussen kan het aandeel TomTom gebruikers in de totale verkeersstroom bepaald worden. Een dergelijke studie naar de ruimtelijke en temporele representativiteit van TomTom data kan uitgevoerd worden door een vergelijking te maken met verkeerstellingsdata die in de Nationale databank wegverkeersgegevens (NDW) beschikbaar is. Binnen de regio Amsterdam betreft dit vooral telpunten op de autosnelwegen en enkele punten op andere hoofdwegen (zie Figuur 5).



Figuur 5 Beschikbare verkeerstelllocaties in de Nationale Databank wegverkeersgegevens (NDW)

3.2. Analyse van reistijdbetrouwbaarheid in de ochtendspits

Om de onderzoeksmogelijkheden van floating car data verder te verkennen hebben we met WVL een analyse opgezet naar reistijdbetrouwbaarheid in de ochtendspits op twee “deur tot deur” trajecten met zowel hoofd- als onderliggend wegennet. Onderzoek naar reistijdbetrouwbaarheid is al wel regelmatig gedaan voor het hoofdwegennet (b.v. Hilbers et al., 2004; Tu, 2010; RWS, 2016), maar veel minder vaak voor het onderliggend wegennet. Floating car data biedt de kans ook dat deel van reistrajecten mee te nemen en daarmee een complete deur tot deur beschrijving te genereren.

Voor dit onderzoek heeft NDW data ter beschikking gesteld voor de maand juni 2017 op basis van gegevens die zij sinds dit jaar van BeMobile geleverd krijgen. Met deze data berekenen we per minuut de totale reistijd per traject bij een verondersteld vertrek tussen 6 en 10 uur in de ochtend. In overleg met WVL zijn twee trajecten gekozen van ongeveer 40 km (enkele) reis die te beschouwen zijn als regulier woon-werkverkeer:

- Westsingel, Oudewater naar Boompjes 200, Rotterdam;
- Middenmolenlaan, Gouda Goverwelle naar Rijnstraat 8, Den Haag.

De herkomst en bestemming van deze trajecten zijn voorgedragen door WVL of in overleg met hen vastgesteld. De route tussen herkomst en bestemming is door NDW bepaald op basis van hun eigen routenetwerk data. Figuur 6 geeft de gekozen routes in meer detail weer, terwijl Tabel 2 de algemene kenmerken van de routes samenvat.



Figuur 6 Overzicht geselecteerde routes

Tabel 2 Algemene kenmerken van de geselecteerde routes

Route	Totale lengte (km)	Lengte hoofd-wegennet (km)	Lengte onderliggend wegennet (km)	Free flow reistijd* (min.)
Oudewater-Rotterdam	41.1	30.6	10.5	27.5
Gouda-Den Haag	34.9	25.8	9.1	24.4

*De free flow reistijd is bepaald als som van de van de 5^e percentiel reistijden van alle afzonderlijke wegvakken waargenomen over de hele periode

De reistijden in deze analyse zijn representatief voor de trips die in deze periode over de gekozen routes gemaakt zouden kunnen zijn, maar ze niet gebaseerd zijn op individuen die dit traject van deur tot deur hebben afgelegd. De reistijden zijn een optelsom van de reistijd over de afzonderlijke segmenten waaruit het traject bestaat, maar die door steeds verschillende individuen zo ervaren is.

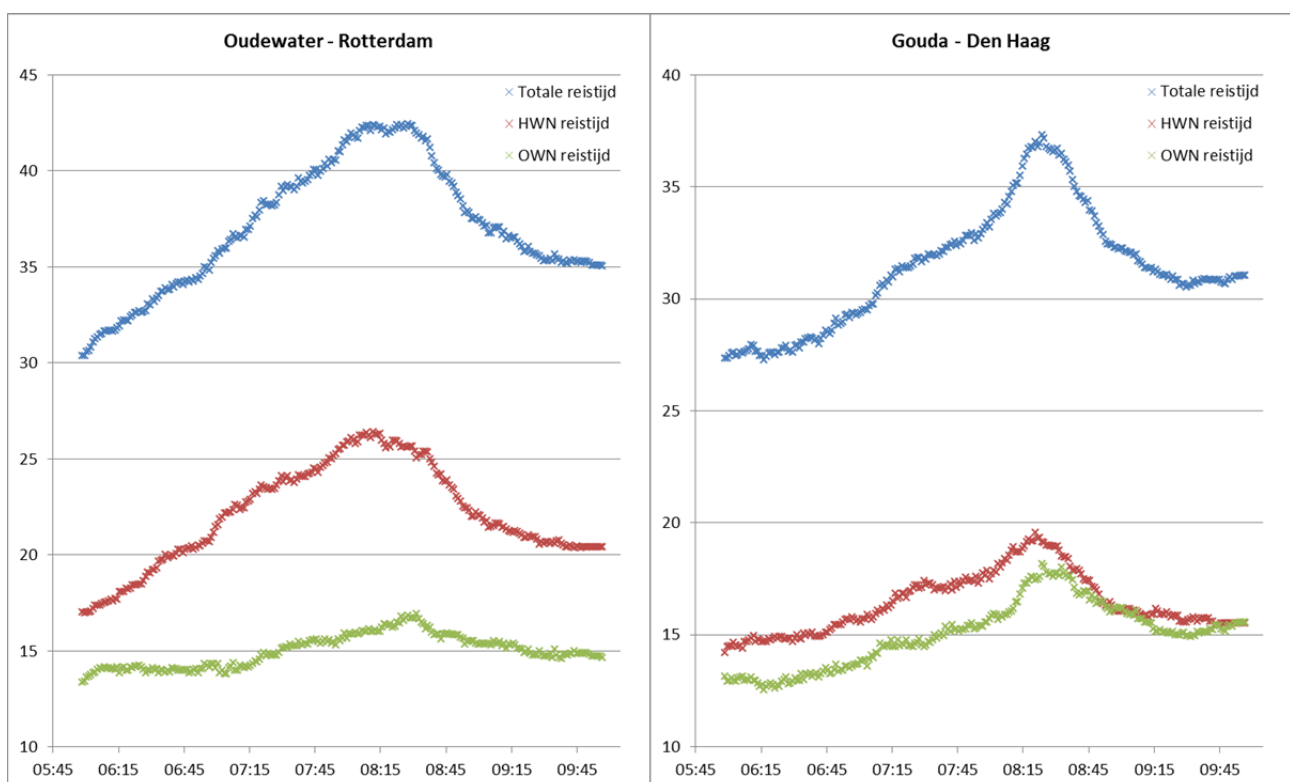
De gekozen routes bevatten een vergelijkbaar lang deel onderliggend wegennet aan het begin (van huisdeur tot snelweg) en eind (van snelweg tot kantoordeur) van de reis. Deze analyse bouwt voort op een recente studie die het bureau Significance uitvoerde in opdracht van WVL (Kouwenhoven en Bel, 2014). In deze studie is voor een beperkte set routes de reistijd geanalyseerd op trajecten met een stuk onderliggend wegennet op basis van een combinatie van lus data (hoofdwegen) en videocameradata met nummerbordherkenning (onderliggend wegennet). Interessant is dat zij geen verband lijken te vinden tussen reistijdbetrouwbaarheid op het hoofdwegennet en onderliggend wegennet (Kouwenhoven en Bel, 2014: p.40). Maar dat kan ook samenhangen met het feit dat zes van de tien door hen gekozen trajecten langer dan 100 km waren, waardoor er een relatief grote spreiding in tijd en ruimte zit tussen de verschillende segmenten.

Onze analyse kijkt op een aantal punten af van de studie van Significance. Allereerst keken zij naar de betrouwbaarheid voor drie dagdelen (ochtendspits, avondspits en de periode daarbuiten), terwijl wij kozen voor de ochtendspits. Daarnaast hebben we kortere routes gekozen die meer overeenstemmen met regulier woon-werkverkeer. Daarnaast was onze geografische focus (Zuidvleugel Randstad) ook beperkter. Het doel van onze studie was ook niet om de studies een op een te kunnen vergelijken voor specifieke locaties (want die data was ook niet direct beschikbaar). Door verschillende manieren te gebruiken om reistijdbetrouwbaarheid uit te drukken is getracht onze uitkomsten beter vergelijkbaar te maken met de studie van Significance. Zoals verderop wordt toegelicht, is afgezien van het berekenen van weekdag-specifieke vertragingen omdat de nu beschikbare tijdreeks van slechts een maand te weinig bruikbare dagen bevatte om dit apart te berekenen.

Gemiddelde verloop ochtendspits

De data die NDW beschikbaar stelde heeft betrekking op de ochtendspitsperiode van 6:00-10:00 op alle dagen in juni 2017 en is uitgeleverd als combinatie van geografische data (ESRI shapefiles met wegsegmenten en hun ID's) en tabel data (kenmerken per routesegment per minuut). Deze data is ingelezen in een open source omgeving (GeoDMS) om reistijden te berekenen. De berekende reistijden houden rekening met het zich tijdens de reis ontwikkelende verkeer. Bij het opbouwen van de trip wordt steeds de reistijd (in milliseconde) gebruikt van de minuut waarin op dat moment gereisd wordt. Dus als de cumulatieve reistijd bij vertrek om 6:00 langer dan 1 minuut is, wordt overgestapt op de reistijd van 6:01. Daarbij wordt rekening gehouden met de locatie waar de voertuigen zich op dat moment bevinden. Na verstrijken van een minuut wordt het resterende deel van de reistijd op dat wegvak gebaseerd op de nieuwe minuut. We gaan er vanuit dat automobilisten maar een kant uit mogen (hier richting Den Haag of Rotterdam) en niet mogen stoppen. Dus er kan niet heen en weer gesprongen worden in de theoretische situatie dat een minuutje wachten op een eerdere link tijds winst oplevert later.

Figuur 7 geeft het gemiddelde verloop van de ochtendspits op werkdagen weer voor beide trajecten. Hiervoor is voor elke minuut de reistijd gemiddeld over alle werkdagen van de maand juni¹³. Daarbij is onderscheid gemaakt in het deel dat over het hoofdwegennet (HWN) wordt afgelegd en het deel dat over het onderliggend wegennet (OWN) verloopt. De reistijd naar Rotterdam is wat langer dan naar Den Haag en bevat ook een relatief langer deel over hoofdwegen. Met name op het traject naar Rotterdam zie je de spits langzaam opbouwen en snel inzakken. Dit is een bekend beeld uit de vervoerseconomie (bottleneck model), forensen willen graag allemaal op dezelfde tijd binnen zijn (in dit geval voor 8:30) en zijn liever te vroeg dan te laat. Verder valt op dat na de spitspiek de free flow snelheden niet meer gehaald worden. Overigens zijn de resultaten in het laatste half uur niet erg betrouwbaar: daar missen we relevante reisweerstanden in de staart van de route (we hebben die immers maar tot 10:00, alles daarna is gekoppeld aan de reisweerstanden van 10:00). Bij vertrek om 9:59 is bijna de hele route is berekend met de weerstanden van 10:00. In de verdere analyses kijken we daarom alleen naar de reizen met vertrek tot 09:30 omdat reizigers dan rond 10:00 op hun bestemming zijn.

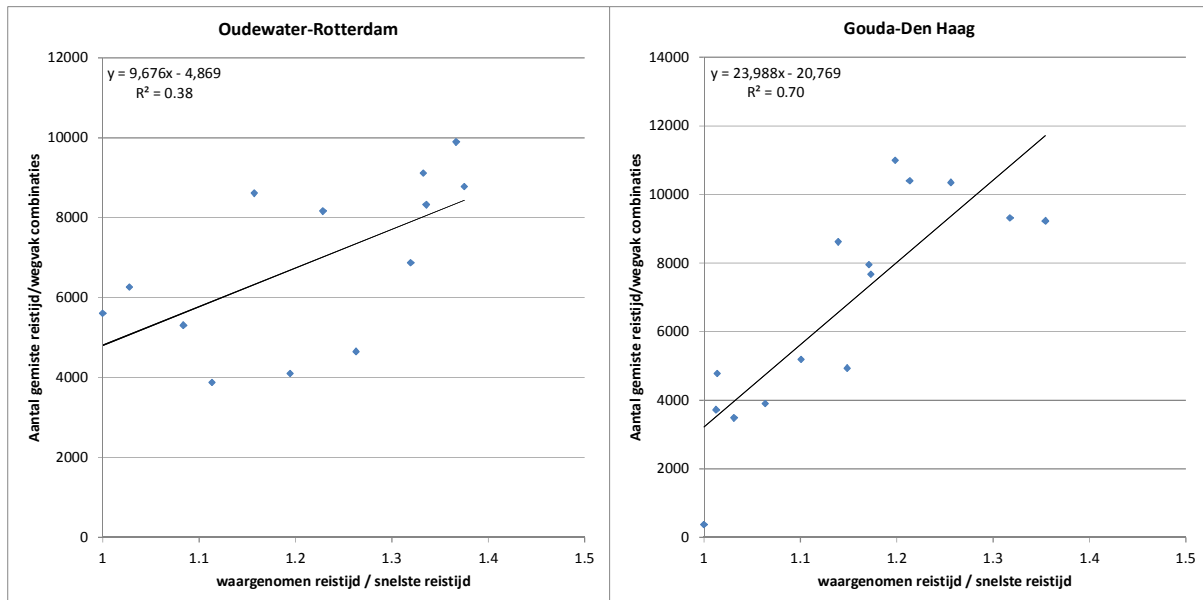


Figuur 7 Gemiddelde reistijd (minuten) tijdens de ochtendspits op werkdagen in juni 2017

Een belangrijk kwaliteitsaspect van de gebruikte data is dat we voor veel segmenten geen metingen hebben. Voor het traject Gouda-Den Haag gaat het hier om 3.5% van de unieke wegsegment-minuut combinaties en voor het traject Oudewater-Rotterdam zelfs om 6.6% (bij vertrek tussen 6:01-9:30 op een van de 20 beschouwde werkdagen). Dit geldt bijvoorbeeld voor de vroege ochtend en dan met name voor het onderliggend wegennet. Hierbij is het goed te bedenken dat wegsegmenten heel kort zijn (gemiddeld 43 meter op het traject naar Den Haag) en deze in het bebouwd gebied vrij kleine wegen kunnen betreffen. Het komt kennelijk geregeld voor dat er binnen een minuut geen voertuig voorbijkomt dat in de BeMobile data geregistreerd wordt. Maar ook tijdens het drukke deel van de spits ontbreken geregeld waarnemingen voor specifieke wegsegment-minuut combinaties. Dit zou

¹³ Van de 22 doordeweekse dagen vallen maandag 5 juni (tweede Pinksterdag) en dinsdag 6 juni af. Gedurende de spits op laatstgenoemde dag is de netwerk shapefile door BeMobile geactualiseerd waardoor een eenduidige opbouw van reistijden ernstig bemoeilijkt wordt.

een gevolg kunnen zijn van de beperkte doorstroming op de wegsegmenten. Als een voertuig meer dan een minuut stilstaat zal op dat moment geen verplaatsing geregistreerd worden. Figuur 8 geeft een overzicht van het aantal gemiste waarnemingen per kwartier in relatie tot de relatieve lengte van de reis. Hieruit wordt duidelijk dat met name bij relatief hoge reistijd (de piek van de spits) er veel waarnemingen gemist worden. Maar ook op andere momenten ontbreken waarnemingen, mogelijk als gevolg van beperkte GPS ontvangst (tunnels?) of GSM-bereik.



Figuur 8 Aantal gemiste waarnemingen per kwartier (tussen 6:01 en 9:30) in relatie tot de relatieve reisduur voor de trajecten Oudewater-Rotterdam (links) en Gouda-Den Haag.

In de gevallen dat een waarneming op een segment ontbreekt gebruiken we de reistijd van het meest nabij liggende tijdstip op hetzelfde segment¹⁴. Deze keuze beïnvloedt natuurlijk de uitkomsten en zal veelal de berekende variantie verlagen: we baseren deze immers voor een klein deel van de wegvakken op gelijk veronderstelde waarden. Maar door de momenten zonder waarnemingen niet te gebruiken zouden we voor veel tijdstippen geen reistijd kunnen berekenen. Een missende waarneming op één wegvak (van gemiddeld 40 meter) betekent dan immers dat we de reistijd over het complete traject (van ongeveer 40 km) niet kunnen bepalen. Alleen reistijden berekenen over compleet beschikbare trajecten is niet mogelijk omdat voor bijna elk vertrekmoment wel een waarneming op een wegsegment ontbreekt. Voor slechts 2.1% van de vertrektijden is data voor het volledige traject beschikbaar.

Overigens past BeMobile ook al een aggregatie over voorliggende minuten toe. Die staat als volgt vermeld in de door NDW meegeleverde data omschrijving: “Be-mobile gebruikt altijd de data van de 10 laatst gepasseerde probes. Dit getal (Coverage) geeft het aantal voertuigen waar de snelheid en reistijd op gebaseerd zijn. De maximale duur waarover data van voertuigen meegenomen wordt is 30 minuten.” De gaten vallen dus pas als er 30 minuten geen probe (= metend voertuig) gepasseerd is. Dit is een van de logische beperkingen van het gebruik van floating car data die is gebaseerd op een beperkte steekproef van voertuigen. Dit betekent dat de gebruikte data op de momenten met

¹⁴ Het komt ook regelmatig voor dat er op bepaalde wegsegmenten de gehele dag geen waarneming gedaan is. In dergelijke gevallen is een free flow snelheid aan het segment toegevoegd die over een langere periode is bepaald. Deze segmenten hebben we echter niet gebruikt in de in deze rapportage gepresenteerde resultaten, waardoor we nu voor 13 (Den Haag) en 9 (Rotterdam) dagen gegevens kunnen gebruiken. De resultaten voor de missende momenten zijn in de geleverde Excel file wel te generen door het filter misslinks = 0 uit te schakelen.

beperkte dekking de daadwerkelijke variatie in reistijd zullen onderschatten. De door ons bepaalde reistijd is kan immers tot een maximum van 30 opeenvolgende minuten zijn gebaseerd op een enkele waarneming. Voor de drukke delen van de spits zou het hergebruiken van waargenomen reistijden in volgende minuten wel tot ingevulde wegvakken moeten leiden, maar dat blijkt in onze analyse niet het geval te zijn. Bij vervolgonderzoek zou in overleg met NDW onderzocht kunnen worden wat de oorzaak is voor ontbreken van waarnemingen op deze en andere momenten en hoe in voorkomende gevallen een betere inschatting van reistijden gecreëerd kan worden.

Reistijdbetrouwbaarheid

Op basis van de berekende reistijden is vervolgens onderzocht wat de reistijdbetrouwbaarheid was op beide trajecten. Reistijdbetrouwbaarheid wordt vaak uitgedrukt in de standaard deviatie ten opzichte van een gemiddelde reistijd (b.v. Peer et al, 2012; Fosgerau, 2010). We hebben er hier voor gekozen de temporele variatie in reistijd te bepalen per tijdvak van 15 minuten en naast de betrouwbaarheid voor het hele traject ook onderscheid te maken in de delen op het hoofd- en onderliggend wegennet. De reistijden zijn beschikbaar voor elke beschikbare vertrekminuut, en de bepaling van variatie per 15 minuten impliceert dat er voor elke dag 14 tijdvakken met waarnemingen beschikbaar zijn tussen 6:01 en 9:30.

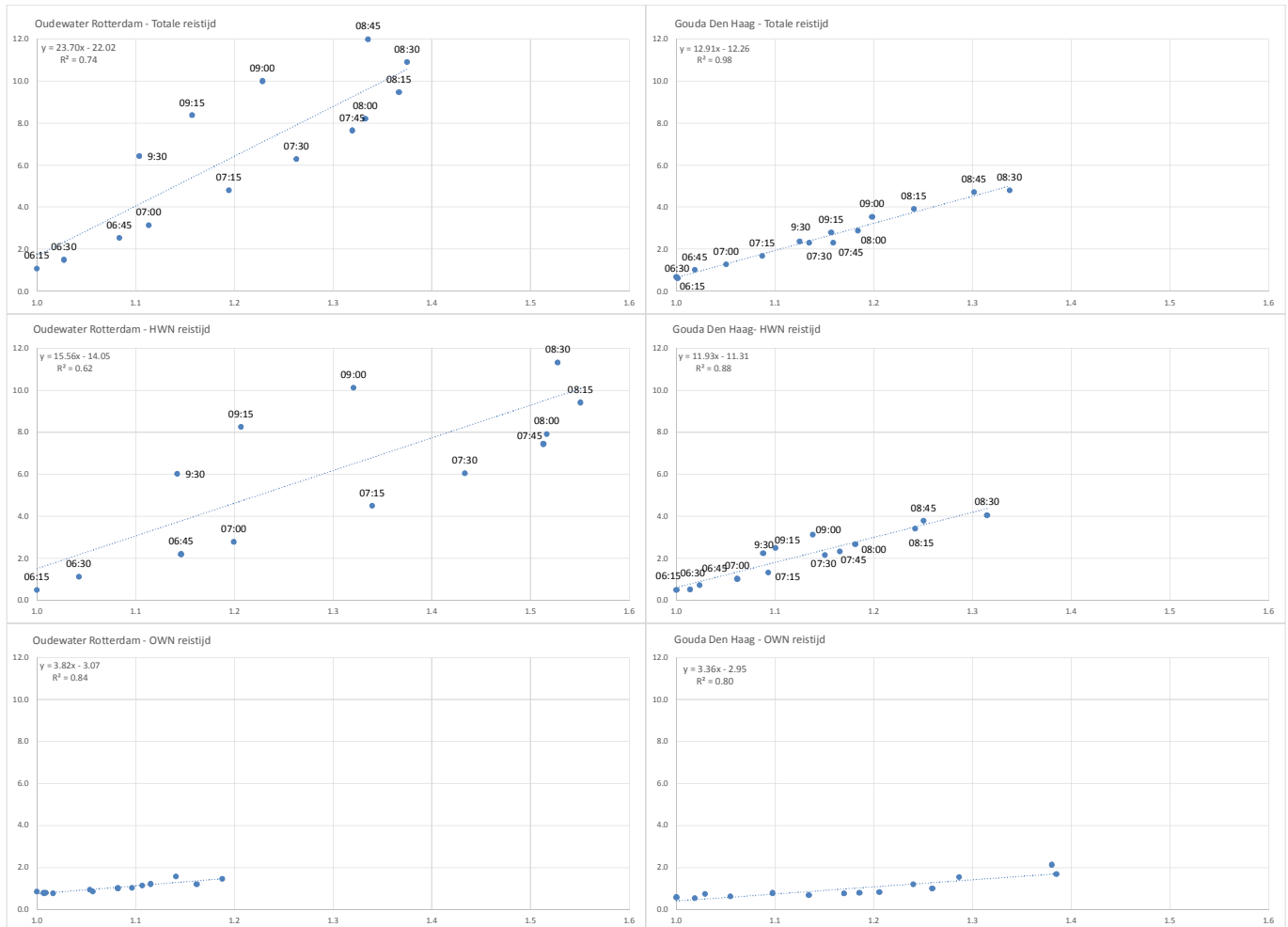
Figuur 9 geeft voor beide trajecten per kwartier de relatieve gemiddelde reistijdvertraging (per traject gedeeld door de kortste reistijd) en bijbehorende standaard deviatie (in minuten) aan. De figuur bevestigt het beeld uit de literatuur dat er een sterke relatie bestaat tussen reistijdvariatie en gemiddelde reistijdvertraging: als op een bepaald tijdstip gemiddeld meer vertraging is, neemt ook de onbetrouwbaarheid toe. Deze relatie is met een lineaire trendlijn in de figuren weergegeven. Bij 10% toename van de reistijd, neemt de standaard deviatie toe met 1.2 minuut richting Den Haag en zelfs 2.3 minuut richting Rotterdam. Het is interessant om te zien dat die onbetrouwbaarheid relatief laag is in de vroege spits en relatief hoog in de na spits. Dat beeld is vooral sterk richting Rotterdam en komt overeen met eerdere studies voor Kopenhagen (Fosgerau, 2010). Duidelijk is ook dat dit voor een belangrijk gedeelte wordt door het hoofdwegennet. De variatie op het onderliggend wegennet is veel geringer. Een belangrijke kanttekening daarbij is dat de variatie hier onderschat kan zijn doordat een groter deel van de reistijden geschat en daarmee minder variabel is. Maar hier speelt ook mee dat de variatie in snelheid geringer is (door de lagere maximum snelheid en bijvoorbeeld de aanwezigheid van verkeerslichten).

De verschillen tussen beide routes kunnen verschillende oorzaken hebben:

- De route naar Den Haag heeft een minder groot verschil tussen minimale en maximale gemiddelde reistijden, en minder momenten dat de gemiddelde totale reistijd meer dan een factor 1.3 is van de laagste gemiddelde reistijd. Mogelijk is op deze route minder sprake van hypercongestie¹⁵, en zijn automobilisten daarom wat minder kritisch op bottleneck-problemen en dus het moment van uitreizen?
- Maar het kan ook zijn dat voor de route naar Den Haag meer (rustige) dagen gebruikt zijn dan voor de route naar Rotterdam, waardoor de gemiddelde reistijd lager ligt.
- Of hebben werkenden in Den Haag (meer overheid) meer flexibele werktijden dan werkenden in Rotterdam (meer private sector).

¹⁵ Het A20 traject tussen Oudewater en Rotterdam bevat twee knelpunten uit de file top-30 van 2017 (Moordrecht-Nieuwerkerk aan de IJssel op nummer 7 en Crooswijk-Rotterdam Centrum op nummer 23) terwijl het A12 traject tussen Gouda en Den Haag pas op positie 36 figureert (Gouwe-Zevenhuizen) en de gebruikelijke file tussen Den Haag-Centrum en Malieveld (nr. 46) net mist. Bron: <https://www.mirtoverzicht.nl/mirt-gebieden/nationaal/file-top-50> (geraadpleegd op 1 december 2017).

- Wat ook mee kan spelen is dat op de route tussen Gouda en Den Haag betrekkelijk weinig verkeer over wegen gaat die ook een lokale functie hebben, terwijl de route naar Rotterdam voor een groot deel over de ring van Rotterdam gaat. Verkeersdrukke op die ring kan betrekkelijk sterk afhangen van fluctuaties in verkeersvraag door lokaal relevante beslissingen - bijvoorbeeld mensen die de fiets laten staan als het regent.

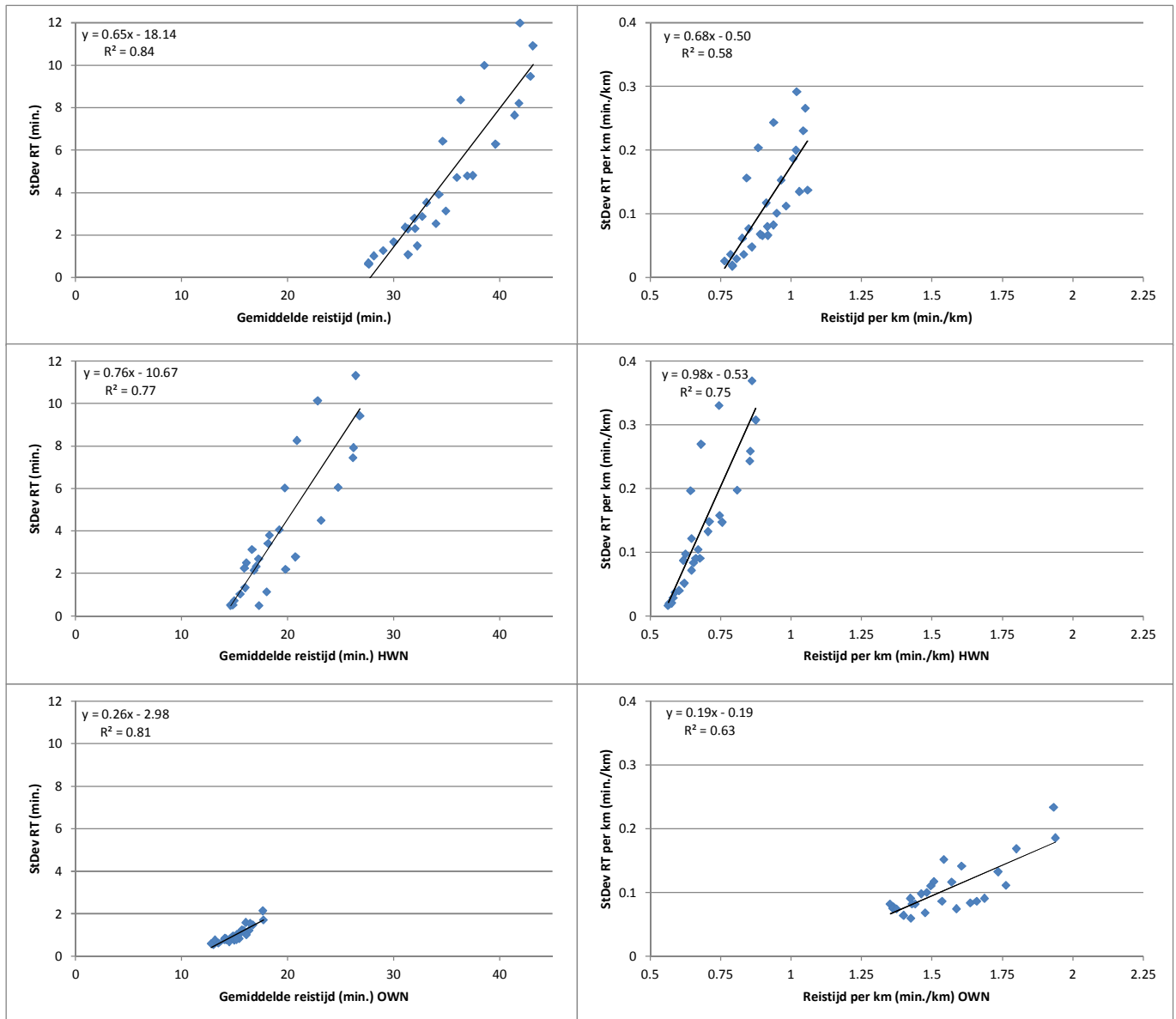


Figuur 9 Variatie in reistijd uitgedrukt in standaard deviatie (minuten) ten opzichte van de gemiddelde reistijd per tijdvak van 15 minuten (per traject gedeeld door de kortste reistijd). De (variatie in) reistijd is bepaald voor het traject als geheel (boven) en afzonderlijk voor het deel op het hoofdwegennet (midden) en het onderliggend wegennet (onder).

Figuur 10 en 11 geven een overzicht van de relatie tussen reistijd onbetrouwbaarheid en verschillende indicatoren van vertraging. Beide trajecten zijn samengenomen in deze figuren, dus elke waarneming in de grafiek komt overeen met een per kwartier en per traject gemiddelde reistijd en bijbehorende standaard deviatie. Voor de twee trajecten en 14 beschouwde tijdvakken zijn dit totaal 28 uren waarnemingen. Door verschillende weergaven van reistijdbetrouwbaarheid in deze figuren op te nemen sluiten we beter aan bij de maten die Kouwenhoven en Bel (2014) in hun studie gebruikten.

In Figuur 10 maken we onderscheid naar gemiddelde reistijd (als in Figuur 9, maar nu niet gestandaardiseerd op de kortste reistijd) en gemiddelde reistijd per kilometer. Als we kijken naar de absolute reistijd (linkerhelft Figuur 10) valt op dat voor de gekozen trajecten de reistijd over het hoofdwegennet gemiddeld langer is, maar ook een grotere variatie en dus onbetrouwbaarheid kent dan de reistijd over het onderliggend wegennet. Dat blijkt zowel uit de grotere range van gemiddelde reistijden per tijdvak van 15 minuten (X-as) als uit de veel grotere standaard deviatie per

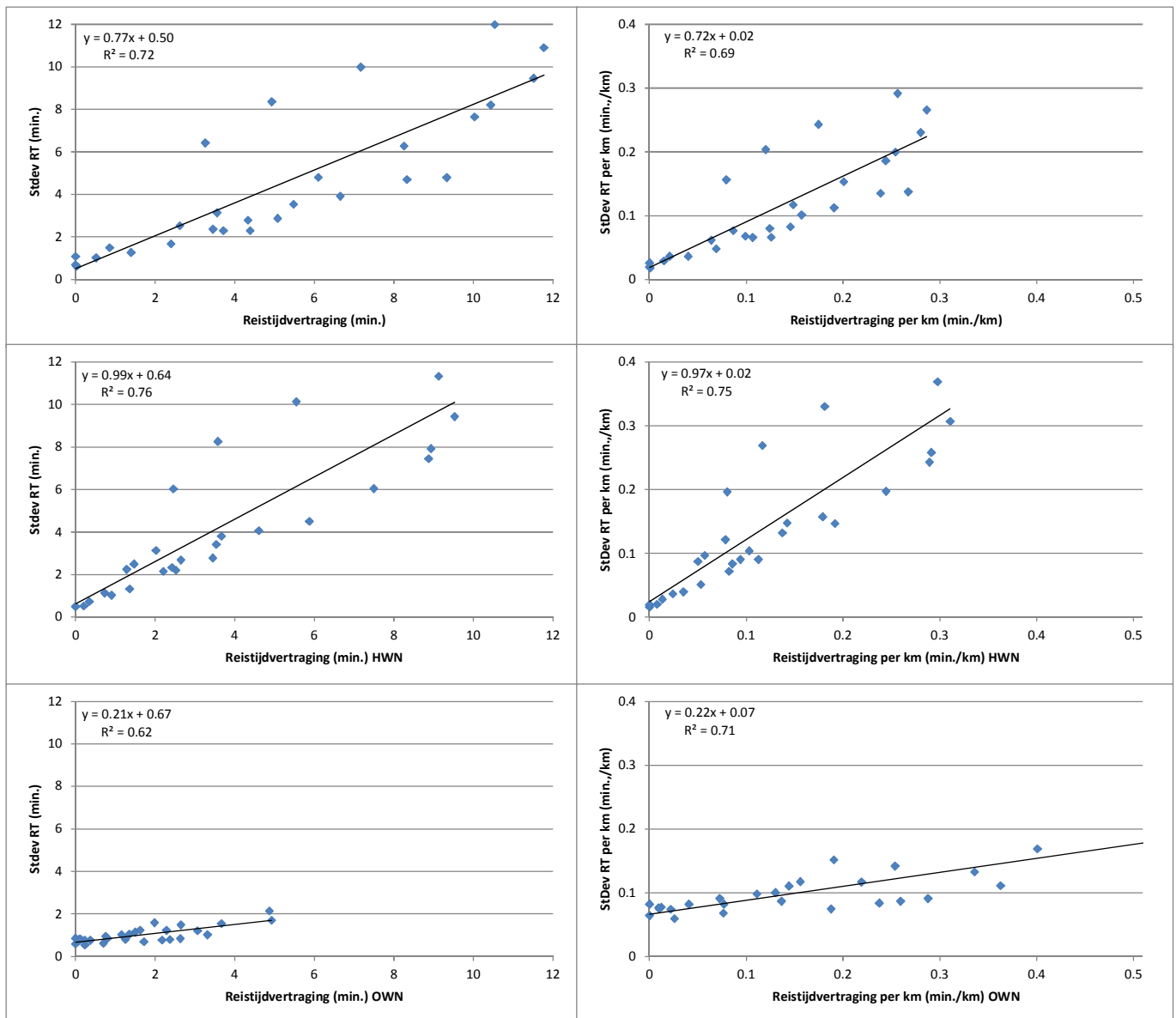
tijdvak (Y-as). In absolute zin draagt het onderliggend weggennet dus maar beperkt bij aan de totale reistijd/onbetrouwbaarheid. De figuur geeft wel aan dat de onbetrouwbaarheid ook hier toeneemt bij langere reistijd. Als we de reistijd normaliseren voor het aantal afgelegde kilometers (rechterhelft Figuur 10) wordt dat beeld nog duidelijker: de standaard deviatie van de reistijd per kilometer kent een vergelijkbare orde van grootte (range op de Y-as) en heeft een duidelijke positieve relatie met toenemende reistijd per kilometer. Uit de lagere coëfficiënten voor de geschatte trendlijnen voor het onderliggend weggennet blijkt verder dat de onbetrouwbaarheid op het onderliggende weggennet minder sterk toeneemt bij langere reistijd dan bij het hoofdwegennet.



Figuur 10 Variatie in reistijd (RT) uitgedrukt in standaard deviatie (minuten) ten opzichte van gemiddelde reistijd (links) en standaard deviatie dichtheid (minuten per kilometer) ten opzichte van gemiddelde reistijd per kilometer (rechts) voor het gehele netwerk (boven), hoofdwegennet (midden) en onderliggend weggennet (onder).

Als naar reistijdvertraging (aantal minuten langzamer dan snelste reistijd op dat traject, in de linkerhelft van Figuur 11) gekeken wordt komt een vergelijkbaar beeld naar voren. Het onderliggend weggennet draagt in absolute zin minder bij aan de reistijdvertraging en de onbetrouwbaarheid is hier geringer dan op het hoofdwegennet. Als de reistijd/onbetrouwbaarheid wordt genormaliseerd naar hoeveelheid kilometers in het traject (rechterhelft van Figuur 11), wordt duidelijk dat dit deel

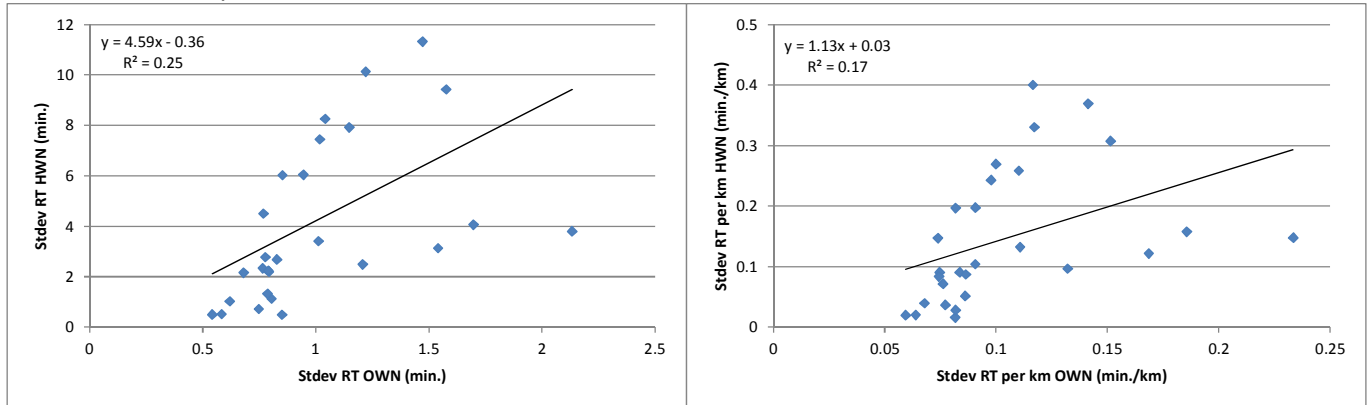
weliswaar minder, maar nog steeds substantieel bijdraagt aan totale reistijd betrouwbaarheid. Bovendien blijkt ook in deze weergave dat de reistijd betrouwbaarheid op het onderliggend wegennet toeneemt bij grotere gemiddelde vertragingen (maar minder sterk dan het hoofdwegennet). Getuige de gevonden R^2 waarden wordt de variantie in absolute reistijdvertraging structureel iets minder goed verklaard op het onderliggend dan op het hoofdwegennet, maar dat verschil is minder sterk als reistijdvertraging gecorrigeerd wordt voor de lengte van het traject dat op onderliggend of hoofdwegennet wordt afgelegd. De beperktere variantie op het onderliggend wegennet zou een gevolg kunnen zijn van de dempende werking van bijvoorbeeld kruisingen en verkeerslichten.



Figuur 11 Variatie in reistijd (RT) uitgedrukt in standaard deviatie (minuten) ten opzichte van gemiddelde reistijdvertraging (links) en variatie in reistijd uitgedrukt in standaard deviatie dichtheid (minuten per kilometer) ten opzichte van de gemiddelde reistijdvertraging per kilometer voor het gehele netwerk (boven), hoofdwegennet (midden) en onderliggend wegennet (onder).

Als laatste hebben we onderzocht in welke mate de reistijdbetrouwbaarheid (uitgedrukt als standaard deviatie) op het onderliggend en hoofdwegennet onderling gerelateerd is. Figuur 12 laat zien dat er een positief verband tussen beide bestaat. Het verband is niet heel sterk (R^2 ongeveer 0.2), wat mede een gevolg is van het feit dat de betrouwbaarheid op beide trajecten (die ook in deze

figuur gecombineerd zijn) verschillend is. Maar dit is desalniettemin een interessant resultaat, dat afwijkt van de studie van Kouwenhoven en Bel (2014) waarin de coëfficiënten voor de correlaties tussen de standaardafwijkingen van de reistijden op de snelwegen en overige wegen voor de 10 onderzochte trajecten varieerden tussen -0.02 en +0.21. Gezien het beperkte aantal trajecten in beide studies kunnen hier geen al te stellige conclusies aan verbonden worden, maar het zou interessant zijn dit nader te onderzoeken voor een groter aantal routes en een langere onderzoeksperiode.



Figuur 12 Relatie tussen variatie in reistijd (RT) op het hoofdwegennet en onderliggend wegennet uitgedrukt in standaard deviatie (minuten, links) en standaard deviatie dichtheid (minuten per kilometer, rechts).

Met de verzamelde gegevens en berekende reistijden kunnen nog verdere analyses gedaan worden. Zo kan gekeken worden naar de reistijden op verschillende weekdays, naar het effect van het toevoegen van de geschatte reistijden voor de wegsegmenten waarvoor geen waarnemingen op de dag zelf gedaan zijn, of juist naar het uitsluiten van reistijden die gebaseerd zijn op nabijgelegen minuten. De apart opgeleverde Excel tabel bevat alle berekende reistijdgegevens en door aanpassing van de filters in de draaitabel kunnen deze gegevens betrekkelijk eenvoudig op andere wijze geaggregeerd worden.

Een belangrijk nadeel van de huidige dataset is echter dat de hoeveelheid gegevens te beperkt is om betrouwbare uitspraken te doen over, bijvoorbeeld, individuele weekdays. Als een langere tijdreeks beschouwd zou worden zouden dat soort analyses gedaan kunnen worden, of zou bijvoorbeeld gekeken kunnen worden naar het effect van specifieke condities zoals weersinvloeden, incidenten of evenementen op reistijd onbetrouwbaarheid. Daarbij zou het natuurlijk interessant zijn ook de avondspits te bezien en meer trajecten te beschouwen om te zien in hoeverre de waargenomen relaties algemener geldend zijn, of bijvoorbeeld sterk afhangen van de filegevoeligheid van (delen van) het traject.

4. Conclusie

Deze rapportage presenteert de resultaten van een beknopte studie naar de toepassing van floating car data voor het beschrijven verplaatsingsgedrag. Het beschrijft de eerste inzichten uit onze literatuurinventarisatie en brengt verslag uit van de eerste exploraties met de data waar TomTom ons toegang toe gaf. Aangezien deze data alleen de waargenomen snelheden en aantallen (TomTom) gebruikers bevat is het niet mogelijk om hier direct herkomst-bestemmingsmatrices mee te schatten zoals eerder voorzien was. In overleg met WVL hebben we dit project daarom afgerond met een verkenning van de mogelijkheden om reistijdbetrouwbaarheid af te leiden uit de floating car data uit een andere bron (BeMobile). Ook deze data beschreef snelheden op individueel wegvak niveau met een hoge temporele resolutie (elke minuut) zonder informatie over de herkomst of bestemming van de geregistreerde voertuigen.

Wegvak en tijdstip specifieke gegevens over snelheden en intensiteiten van een selectie van weggebruikers is ook bruikbaar voor andere onderzoeksonderwerpen dan zijn besproken in dit rapport. Deze onderwerpen presenteren we kort in dit afsluitende hoofdstuk. In de kern richten onze suggesties zich op het analyseren van de ruimtelijke aspecten van de data door de tijd heen, op een temporeel aggregatie niveau dat aansluit op verkeersmodellen die nu door het RWS worden ingezet. Een beter inzicht in de structurele ruimte-tijd dynamiek in verkeersstromen kan gebruikt worden om herkomst-bestemmingsmatrices te schalen, en zo relatief eenvoudig de bestaande verkeersmodellen aan te passen om beter representatief te zijn voor bepaalde momenten in het jaar.

In de afgelopen paar jaar heeft de nationale databank wegverkeersgegevens (NDW) enkele eerste verkenningen gedaan heeft naar de bruikbaarheid van floating car data. We presenteren daarom eerst een korte samenvatting van twee onderzoeksrapporten die door het NDW zijn opgesteld. In aanvulling daarop beschrijven wij een aantal aanvullende onderzoeksopties gebaseerd op onze eigen verkenning van deze data. Als laatste reflecteren we ook nog kort op de mogelijkheden die geboden worden door alternatieve floating-car data bronnen.

4.1. Recent onderzoek naar bruikbaarheid floating car data per weg segment

Recent heeft NDW twee pilots uitgevoerd waarin floating car data (FCD) is vergeleken met de gangbare data van vaste meetlocaties die NDW ontsluit. NDW is zo vriendelijk geweest ons de rapporten van beide studies ter beschikking te stellen. Het betreft:

- FCD Pilot Zuid Holland. Niet openbaar onderzoeksrapport met vergelijking tussen drie floating car data leveranciers en NDW-data in de provincie Zuid-Holland. Opgesteld in 2016 door Marthe Uenk-Telgen, Bram Bakker en Bjorn Heijligers;
- Datafusie Pilot; Openbare resultaten. Eindrapport R11448. Opgeleverd oktober 2014 door Isabel Wilmink, Diana Vonk Noordegraaf, Tom Thomas, Victor Westerwoudt en Manus Barten.

Beide rapporten vatten we hieronder kort samen. Daarnaast deed NDW recent een vervolgonderzoek waarin de analyses van Zuid-Holland uitgevoerd zijn op trajecten in de gemeente Amsterdam, Groningen, Leeuwarden en de provincie Overijssel. De resultaten laten we hier buiten beschouwing.

Onderzoeksrapport FCD Pilot Zuid Holland

Het rapport richt zich op floating car data van drie verschillende providers, en behandelt uitsluitend real-time datafeeds met 5 minuut-interval voor 50 trajecten in de provincie Zuid-Holland. Belangrijk hierbij is het omgaan met de stochastische binnen dit frequentiedomein en de bruikbaarheid van

gefilterde data voor het tijdig signaleren van verkeersincidenten en congestie met een lage foutmarge.

Samengevat had het onderzoek de volgende doelstellingen:

- kwaliteitsonderzoek om te bepalen of fysieke meetpunten voor reistijdinformatie verwijderd kunnen worden;
- kwaliteitsonderzoek om te bepalen of de reistijdinformatie kan worden gebruikt voor het aansturen van DRIPs;
- eerste indruk krijgen van de variatie in kwaliteit tussen leveranciers;
- een “best-practice” opzet ontwikkelen waarmee floating car data geëvalueerd en vergeleken kan worden.

Het rapport heeft dus een duidelijke focus en “...overige vragen, bijv. ten aanzien van verkeersintensiteiten, ..., HB-matrices, etc., vallen hierbuiten..”.

De hoofdconclusie van het rapport luidt: *“De resultaten van de analyse van deze pilot tonen aan dat floating car data geschikt is als bron voor informatie over real-time snelheden/reistijden en om begin en eind en ernst van congestie te detecteren, met voldoende tijdigheid en grote beschikbaarheid en dekking.”*

Enkele belangrijke bevindingen uit het rapport zijn verder:

- de dekkingsgraad van floating car data is hoog;
- op basis van een aantal statistische vergelijkingsmethoden blijkt de snelheid data vergelijkbaar met NDW data, maar niet veel beter dan base-line free-flow reistijd; na toepassen van een convolutie filter op geïnterpoleerde data is hoog frequente floating car data “goed” vergelijkbaar met huidige NDW-systemen;
- er is geen structurele overschatting of onderschatting van snelheden;
- floating car data is bruikbaar voor tijdige signalering van congestieverschijnselen;
- de beschikbaarheid van floating car data in het hoge tijdsfrequentiedomein (minuten) is hoger dan NDW data; dit kan te maken hebben met post-processing aan de kant van de data-provider;
- de resultaten laten zien dat meestal de floating car data iets achterloopt bij de NDW-data, 1 tot een enkele keer wel 10 minuten.

Datafusie Pilot - Openbare resultaten

Het rapport richt zich op het combineren van floating car data van verschillende providers en data van fysieke meetpunten, en behandelt uitsluitend real-time datafeeds met 1 minuut-interval voor het hoofdwegennet en onderliggend wegennet rondom Delft. Ook in deze eerdere studie lag de nadruk op het omgaan met de stochasticiteit en de bruikbaarheid van gefilterde data voor het tijdig signaleren van verkeersincidenten en congestie.

De hoofddoelstelling is het uitvoeren van een peiling om inzicht te krijgen in de mogelijke kwaliteit van wegverkeersgegevens (puntsnelheid, intensiteit en reistijd) uit gecombineerde floating car data, gegevens uit vaste meetlocaties en datafusie bij lagere dichtheden van de vaste meetlocaties.

De hoofdconclusies zijn:

- floating car data en datafusie kunnen goed ingezet worden voor het genereren van gegevens over snelheden;
- de inzet voor het bepalen van intensiteiten is lastiger en hiervoor zijn nog vaste locaties nodig;

- patronen zoals bij filegolven en incidenten lijken doorgaans goed in beeld gebracht te worden;
- een meerwaarde van floating car data en datafusie is dat de visualisaties van deze data een completer verkeersbeeld opleveren;
- uit de vergelijking met de lusdata blijkt niet dat het vervangen van lusdata op het niveau van de huidige NDW kwaliteitseisen, met floating car data en datafusie haalbaar is.

Uit het document wordt niet duidelijk of alleen ruwe data is vergeleken, of dat ook statistische technieken zijn toegepast om de achterliggende tijdreeksen uit de data te filteren.

4.2. Mogelijk vervolgonderzoek met floating car data op wegsegment niveau

Gezamenlijk geven de beide NDW rapporten een duidelijk beeld van de karakteristieken van floating car data met een hoge tijdsfrequentie voor gebruik in (near) real time toepassingen. De meeste bevindingen zijn, echter, niet direct vertaalbaar naar analyses die zich richten op een ander temporeel aggregatieniveau en bijvoorbeeld verplaatsingsgedrag over een langere periode beschouwen.

We noemen hier enkele suggesties voor vervolgonderzoek dat zich richt op het beschrijven en begrijpen van verplaatsingsgedrag. We richten ons specifiek op floating car data die gekoppeld is aan individuele wegsegmenten zoals die wij in deze verkenning tot onze beschikking hadden. De mogelijkheden voor het analyseren van het keuzegedrag van individuen laten we hierbij dus buiten beschouwing omdat daar floating car data voor nodig is waarover RWS momenteel niet beschikt. Bij de onderstaande beschrijving van onderzoeksideeën is een beknopte toelichting opgenomen waarin wordt aangegeven hoe het voorgestelde onderzoek zich verhoudt tot hetgeen dat al door NDW is uitgevoerd. In aanvulling op mogelijk onderzoek naar verplaatsingsgedrag hebben we als laatste ook een meer methodische onderzoekssuggestie toegevoegd die zich richt op enkele belangrijke aspecten van tijdreeksanalyse van data met een hoge temporele resolutie die in het tot nu toe door NDW uitgevoerde onderzoek niet aan de orde zijn gekomen.

Verdiepen reistijdbetrouwbaarheidsanalyse

In onze analyse van reistijdbetrouwbaarheid bleken de gebruikte data en methoden bruikbaar om variatie in reistijd gedurende de spits te beschrijven. Voor de gekozen twee trajecten leverde dat een eerste verkenning op van de betrouwbaarheid van deur-tot-deur reizen in de spits met daarbij inzicht in de verschillen tussen hoofd- en onderliggend wegennet. Deze analyse bevestigt allereerst een bekend beeld: de onbetrouwbaarheid (uitgedrukt als standaard deviatie) wordt groter met toenemende reistijdvertraging: in de piek van de spits zijn reizen niet alleen gemiddeld langer, maar ook nog eens onzekerder. Die verschillen zijn op het hoofdwegennet veel geprononceerder dan op het onderliggend wegennet, maar vertonen wel hetzelfde basispatroon van toenemende onbetrouwbaarheid bij langere reistijden.

Onze bevindingen lijken af te wijken van een van de weinige andere studies over dit onderwerp waarin geconstateerd werd dat “dat de standaardafwijking van de reistijden op het snelwegdeel niet of nauwelijks gecorreleerd was met de standaardafwijking van de reistijden op de overige wegen” (Kouwenhoven en Bel, 2014: p.40). Beide studies gebruiken echter een kleine steekproef (wij slechts 2 trajecten, tegenover 10 bij Kouwenhoven en Bel), heel andere trajecten (circa 40 km lengte in ons geval, tegenover 100 km in de andere studie), andere regio’s (Den Haag/Rotterdam versus een groot deel Nederland), andere perioden (ochtendspits versus drie dagdelen) en wellicht ook een andere wegingdeling (hoofd- en onderliggend wegennet tegenover snelwegen en overige wegen). Daarnaast wijkt ook de methode waarmee onbetrouwbaarheid is uitgedrukt af. Dat betreft zowel de grondslag voor betrouwbaarheid (de gemiddelde reistijd per tijdvak van 15 minuten voor alle beschikbare dagen, tegenover een weekdag specifieke reistijd) als de manier waarop ze gecombineerd (als een

doorlopende reis of als los opgetelde, gescheiden onderdelen). Daarbij worden de standaardafwijkingen op hoofd- en onderliggend wegennet anders vergeleken (op basis van een vergelijkbare relatie ten opzichte van totale reistijd of middels een direct vergelijk van de gevonden standaard afwijkingen per kwartier).

Met de nu bij RWS beschikbare floating car data en de in dit onderzoek toegepaste methode van reistijdanalyse is het redelijk eenvoudig de analyse uit te breiden tot een groter aantal trajecten. Hiermee kan onderzocht worden in hoeverre de waargenomen relaties algemener geldend zijn en wordt tevens een robuuster vergelijk met de studie van Kouwenhoven en Bel mogelijk. Door ook gebruik te maken van een langere onderzoeksperiode wordt het mogelijke weekdagspecifieke analyses te doen en daarmee de vergelijkbaarheid verder te vergroten. Een langere tijdreeks maakt het ook mogelijk analyses te doen naar het effect van specifieke condities zoals weersinvloeden, incidenten of evenementen op reistijdonderbrouwbaarheid. Daarbij is het interessant de reistijden tijdens de avondspits en de rest van de dag te onderzoeken.

Voor een eventueel vervolgonderzoek zou het goed zijn nader in te gaan op mogelijke oorzaken voor deze reistijdinformatie die in deze studie voor een klein deel van de wegvakken ontbrak. Dit zou in overleg met NDW onderzocht kunnen worden en met hen zou ook nagedacht kunnen worden over methoden waarmee in voorkomende gevallen een betere invulling van missende reistijden gerealiseerd kan worden. Voor de begin- en eindpunten van de te selecteren routes op het onderliggend wegennet is het belangrijk die wegvakken te selecteren die een redelijke hoeveelheid verkeer bevatten. Idealiter worden hierbij logische routealternatieven voor dezelfde herkomst-bestemmingscombinaties in samenhang beschouwd. Op basis hiervan kan steeds de snelste route geselecteerd worden, of bijvoorbeeld onderzocht worden in welke mate de snelheid en betrouwbaarheid van alternatieve opties in de tijd verschilt.

Ruimtelijke en temporele dynamiek in verplaatsingsgedrag

De sterkste kant van de data wegsegment data die nu beschikbaar is, is de hoge ruimtelijke en temporele resolutie in waargenomen snelheid. Hiermee kan variatie in snelheid door de tijd geanalyseerd worden. De nadruk hierbij zal dan liggen op het frequentiedomein dat van belang is voor een verkeersmodel. Zo kan er gekeken worden naar de verkeersintensiteit van metingen op het hoofdnetwerk of het onderliggende wegennetwerk. Zijn er structurele verschillen in werkdagen? Zijn er structurele seizoenseffecten? Is die structurele tijdsdynamiek verschillend voor grootstedelijke gebieden t.o.v. dunbevolkte gebieden? Kan de mogelijke ruimtelijk heterogene tijdsdynamiek gebruikt worden om H-B matrixen te schalen, zodat er in verkeersmodellen scenario's uitgedraaid kunnen worden die relevant zijn voor bepaalde momenten in de week of in het jaar?

De NDW-rapporten bevestigen de hoge ruimtelijke en temporele resolutie in waargenomen snelheid. Er is een vergelijking gedaan tussen hoogfrequente databronnen, maar aspecten van de data op een temporeel aggregatieniveau dat aansluit op bestaande verkeersmodellen zijn nog niet benoemd. Er wordt wel genoemd dat stochastische op deze frequentie erg belangrijk is. Bestaande methoden uit de tijdreeksanalyse om met deze stochastische om te gaan zijn echter niet toegepast. Een structurele decompositie van de data is op zowel hoge frequentie als lage frequentie lijkt nog niet uitgevoerd. Een aantal vragen blijft daarmee onbeantwoord. Het lijkt ons dat met name de structurele tijdsdynamiek interessant is voor verkeersmodellen wat impliceert dat vooral naar een geaggregeerd tijdsniveau gekeken zou moeten worden.

Ruimtelijke en temporele representativiteit van de weg-segment data

Een ander belangrijk onderwerp is de ruimtelijke en temporele representativiteit van floating car data. Varieert de dekking van deze data op het hoofdnetwerk of het onderliggende wegennetwerk met verschillende werkdagen? Zijn er structurele seizoenseffecten? Is die structurele

tijdsdynamiek in het marktaandeel verschillend voor grootstedelijke gebieden t.o.v. dunbevolkte gebieden?

Inzicht daarin kan verkregen worden door een vergelijking uit te voeren met de door NDW beheerde verkeerslusdata die alle passerende voertuigen telt. Dat brengt de ruimtelijke en temporele verspreiding van het “marktaandeel” van de leveranciers van floating car data in beeld. Uit het uitgevoerde NDW onderzoek blijkt dat verkeersintensiteit lastig uit floating car data te herleiden valt, maar specifieke cijfers worden daar niet bij genoemd. Beide onderzoeken richten zich daarbij op afgebakende gebieden. Het valt uit het uitgevoerde onderzoek niet te herleiden of floating car data representatief is door heel Nederland, en hoe de representativiteit door de tijd heen varieert.

Vergelijk waargenomen verkeersgedrag met verkeersmodelsimulaties

Het NDW onderzoek bevestigde dat floating car data het voordeel hebben van een hogere dekkingsgraad van het onderliggend wegennetwerk. De snelheidsdata lijkt op hoge frequentie representatief, de intensiteitsdata van floating car data lijkt echter ontoereikend. Door de hoge dekkingsgraad is het zeker mogelijk om goede vergelijkingen te maken tussen de verkeerssnelheidsoutput van de verkeersmodellen die nu door RWS gebruikt worden en waargenomen floating car data gegevens op plekken waar het NDW geen metingen heeft. Interessante vragen daarbij zijn: Hoe verhouden verkeerseigenschappen uit floating car data zich tot de output van door het RWS ingezette verkeersmodellen? Wat lijkt er uit vergelijkingen te komen op wegen waar het NDW geen fysieke meetlocaties heeft?

Verder blijft het een open vraag hoe bruikbaar gestandaardiseerde intensiteitsmaten zijn die gebaseerd zijn op temporele aggregatie. Het is aannemelijk dat voor veel gebieden een vergelijking tussen gestandaardiseerde intensiteitsmaten mogelijk is.

Relatie snelheid en intensiteit

De snelheidsdata lijkt op hoge frequentie representatief. Gestandaardiseerde intensiteitsmaten op basis van een tijdsaggregatie zijn nog niet onderzocht. Zodra er inzicht is in de representativiteit van een gestandaardiseerde intensiteit uit floating car data, kan de hoge dekkingsgraad gebruikt worden om relaties tussen snelheid en intensiteit voor verschillende typen wegen nauwkeurig te beschrijven. Daarbij kan gekeken worden naar vragen als: Hoe verhouden snelheid en intensiteit zich tot elkaar op verschillende typen wegen, en zijn er spillover effecten in snelheden tussen naburige wegsegmenten?

Er kan verder ook gekeken worden hoe de snelheden op een segment zich verhouden tot de intensiteit op naburige onderliggende wegennetwerk. Spillover effecten zijn hierbij van belang, maar ook de volgorde van causaliteit door de tijd heen.

Verbeteren dynamische verkeersinformatie

In het uitgevoerde onderzoek kwam duidelijk naar voren dat de floating car data zeer stochastisch is, en dat het huidige onderzoek een sterke nadruk heeft op hoogfrequente doeleinden. De stochasticiteit in dit frequentiedomein heeft te maken met de soms lage en sterk variërende observatiedichtheid. De gebruikte statistische vergelijkingsmethoden van de ruwe data uit het eerste rapport kunnen in twijfel getrokken worden, met name de conclusie dat de data niet veel beter is dan free flow reistijd komt hierdoor in het geding. Een statistische vergelijking van variabelen met sterke meetfouten is theoretisch ongegrond, aangezien er voor dergelijke variabelen geen consistentietheorie toepasbaar is. In het eerste rapport wordt wel een simpele convolutie filter toegepast, maar dergelijke filters gebruiken gegevens uit de toekomst en kunnen real-time niet toegepast worden. Een betere aanpak, die wel real-time toegepast kan worden, zou zijn om met behulp van state-space modellen, e.g., Kalman filters, de staat van de latente snelheidsvariabele te

schatten en die te scheiden van de meetfouten. Vervolgens kan er een vergelijking uitgevoerd worden op basis van de latente variabele. Een mogelijk interessant onderzoek domein is het voorspellen en filteren van hoog frequente floating car data op basis van econometrische modellen die bijvoorbeeld ingezet worden om hoog frequente financiële tijdreeksen te analyseren.

Hiermee kan antwoord gegeven worden op vragen als: Is het mogelijk om DRIP's voortijdig van informatie te verschaffen, door het gebruik van gefilterde hoog frequente floating car data, waarbij alleen gegevens uit het verleden gebruik worden om tot een gefilterde staat van de snelheidsgegevens te komen? Valt hoogfrequente floating car data te voorspellen met geavanceerde econometrische modellen? Is het mogelijk om real-time trend-breaks te detecteren in real-time gefilterde data?

4.3. Alternatieve floating car data

Naast de floating car data van TomTom en BeMobile die we voor deze studie gebruikten, zijn andere aanbieders (b.v. HERE en Google) actief met vergelijkbare data. Meer informatie over de karakteristieken van deze data (toegankelijkheid, kosten, representativiteit etc.) is te vinden in een recente studie voor (toen nog) het Ministerie van Infrastructuur en Milieu en de Metropoolregio Amsterdam van Goudappel Coffeng (2017). Deze studie beschrijft tevens vrij summier een methode om deur tot deur reistijden voor verschillende transportmodaliteiten te achterhalen. Voor reistijden per auto wordt gebruik gemaakt van onder meer de Google API (interface om actuele reistijden op te vragen) en floating car data van HERE.

Een belangrijk aspect van de nieuwste generatie in-car navigation software is dat deze ook diensten levert die van belang zijn voor de dagelijkse weggebruikers die hun reisroute al kennen. Door real-time informatie over bijvoorbeeld files of flitsers aan te bieden zullen autobestuurders deze applicaties continu gebruiken, wat een veel bredere dekking van het totaal aantal weggebruikers zal geven. Een andere interessante databron betreft het verplaatsingsgedrag van specifieke groepen weggebruikers die apart worden opgeslagen (taxi's, lease auto's wordt gebruikt).

In deze studie is gebruik gemaakt van de floating car data die is opgeslagen per wegsegment. Hoewel deze data bruikbaar is voor diverse toepassingen zoals hiervoor is geïllustreerd, ligt de echte kracht van dergelijke data in het beschrijven van individueel reisgedrag. Dergelijke informatie is bij de verschillende aanbieders van deze data beschikbaar, maar wordt in basis niet uitgeleverd. Dat ligt allereerst natuurlijk aan de privacy gevoeligheid van deze data. Maar daarnaast spelen ook andere aspecten zoals commerciële belangen en de enorme omvang van deze data. Voor het schatten van herkomst-bestemmingsmatrices en andere aspecten van verplaatsingsgedrag (zoals route en tijdstip keuze) is het dus interessant om te bezien of ook trip-data via deze floating car data leveranciers verkregen kan worden.

HERE Traffic Analytics ontwikkelt een interessant product waarmee herkomst-bestemmingsmatrices opgesteld kunnen worden op basis van het reisgedrag van hun gebruikers. Op basis van het verplaatsingsgedrag voor de afgelopen drie jaar kan hiermee voor door de gebruiker op te geven grids of zones het aantal trips tussen deze gebieden opleveren. Uiteraard gelden hier vanuit privacy-oogpunt minimumeisen aan de grootte van deze zones (minimaal 7 trips per zone, en minimaal een aantal inwoners per zone). Dit product vertaalt aantallen HERE gebruikers naar totale verkeersvolumes. Het zou voor RWS zeer interessant zijn de kwaliteit van dit product nader te onderzoeken op bruikbaarheid voor haar modelspecificatie. Vanuit de VU willen we een samenwerking opstarten met HERE om de onderzoeksmogelijkheden van hun data bronnen te verkennen.

Referenties

- Ásmundsdóttir, R. (2008) Dynamic OD matrix estimation using floating car data. MSc thesis – Civil Engineering Delft University of Technology. 148 pp.
- Bera, S., Krishna Rao, K. V. (2011) Estimation of origin-destination matrix from traffic counts: the state of the art. *European Transport / Trasporti Europei* 49: 3-23
- CBS en RWS (2016) Onderzoek Verplaatsingen in Nederland 2015 - OVIN 2015. Centraal Bureau voor de Statistiek en Rijkswaterstaat. <https://doi.org/10.17026/dans-z38-prz4>
- Calabrese, F., Di Lorenzo, G., Liu, L., Ratti, C. (2011) Estimating Origin-Destination Flows Using Mobile Phone Location Data. *IEEE Pervasive Computing* 10 (4): 36-44. DOI 10.1109/mprv.2011.41
- Fosgerau, M. (2010) On the relation between the mean and variance of delay in dynamic queues with random capacity and demand, *Journal of Economic Dynamics and Control* 34 (4): 598-603.
- Goudappel Coffeng (2017) Deur-tot-deur methode; reistijd, betrouwbaarheid, beleving. MIRT onderzoek stedelijke bereikbaarheid MRA. Goudappel Coffeng rapport MI2005/Mqt/0077.02, in opdracht van Ministerie van Infrastructuur en Milieu en Metropoolregio Amsterdam. 53 pp.
- Hilbers, H., Ritsema Van Eck, J., Snellen, D. (2004). Behalve de dagelijkse files; over betrouwbaarheid van reistijd. Den Haag: Ruimtelijk Planbureau. 97 pp.
- Hazelton, M. L. (2003) Some comments on origin–destination matrix estimation. *Transportation Research Part A* 37: 811–822.
- Kouwenhoven, M., Bel, N. (2014) Raming omvang betrouwbaarheid Nederlandse wegennet. Eindrapport (versie 5). Rapport voor Rijkswaterstaat Water, Verkeer en Leefomgeving. Significance. 124 pp.
- Peer, S., Koopmans, C.C., Verhoef, E.T. (2012) Prediction of travel time variability for cost-benefit analysis, *Transportation Research Part A: Policy and Practice* 46 (1): 79-90.
- Peterson, A. (2007) The Origin–Destination Matrix Estimation Problem - Analysis and Computations. Linköping Studies in Science and Technology. PhD thesis. Norrköping, Sweden. 48 pp.
- Reinthal, M., Zajicek, J. (2009) Integration of traffic flow relations in a simulation environment. *Urban Transport XV* 117. doi:10.2495/UT090121
- RWS (2016) Publieksrapportage Rijkswegennet 3e periode 2015. Rijkswaterstaat, Informatiepunt Water, Verkeer en Leefomgeving.
- Simoni, M. (2014) State-of-the-art of the OD-matrix calibration (Inventarisatie Kalibratiemethoden). Rapport in opdracht van Rijkswaterstaat-Ministerie van Infrastructuur en Milieu. TU-Delft. 56 pp.
- Tu, H. (2010) Reistijdbetrouwbaarheidsmodel op autosnelweg. *Verkeerskunde* 2010, 2. <http://www.verkeerskunde.nl/integrale-artikelen/reistijdbetrouwbaarheidsmodel-op-autosnelweg.20137.lynkx>
- Uenk-Telgen, M., Bakker, B., Heijligers, B. (2016) FCD Pilot Zuid Holland. Niet openbaar onderzoeksrapport met vergelijking tussen drie FCD-leveranciers en NDW-data in de provincie Zuid-Holland. Nationale Databank Wegverkeersgegevens. 62 pp.
- Van der Zijpp, N.J. (1997) Dynamic origin-destination matrix estimation from traffic counts and automated vehicle identification data. *Transportation research record* 1607: 87-94. DOI: 10.3141/1607-13
- Van Vuren, T., Carey, C., MacDonald, M. (2011) Building practical origin-destination (OD/trip) matrices from automatically collected GPS data. *Proceedings European Transport Conference 2011*. Association for European Transport.
- Wilmink, I., Vonk Noordegraaf, D., Thomas, T., Westerwoudt, V. Barten, M. (2014) Datafusie Pilot; Openbare resultaten. Eindrapport R11448. DITCM Innovations. 46 pp.